



kwiteit
forensische
zorg

2020

Validatie FARE

Een multi-center studie naar de interbeoordelaarsbetrouwbaarheid, convergente validiteit, veranderbaarheid en de predictieve validiteit van de Forensische Ambulante Risico Evaluatie (FARE)

M.J. Eisenberg, MSc, J.C. Hutten, MSc, dr. F.C.A. van der Veeke, dr. Y.H.A. Bouman, Prof. dr. S. Bogaerts en dr. J.E. van Horn

Datum van publicatie: september 2020

Colofon

Deze publicatie/ dit project is tot stand gekomen met financiële ondersteuning van het Programma KFZ. Het Programma KFZ heeft daarnaast zorg gedragen voor de inhoudelijke projectbegeleiding waarbij is gestuurd op kwaliteit, voorgangsbewaking en budgetbeheer. Het Expertisecentrum Forensische Psychiatrie voert namens het programma KFZ het projectmanagement.

© KFZ: Het intellectuele eigendom ligt bij de auteur(s). De auteur(s) stemt er mee in dat deze uitgave onvoorwaardelijk en zonder kosten gebruikt kan worden door alle instellingen binnen het forensische zorgveld



Projectuitvoering

Dr. J.E. van Horn, hoofd afdeling onderzoek de Waag, onderdeel van De Forensische Zorgspecialisten.

Prof. dr. S. Bogaerts, hoogleraar forensische psychologie, departement ontwikkelingspsychologie, Tilburg University; Directeur-hoogleraar Fivoor Wetenschap en behandelinnovatie.

M.J. Eisenberg, MSc., onderzoeker de Waag, onderdeel van De Forensische Zorgspecialisten en promovenda Universiteit van Amsterdam, Faculteit der Maatschappij- en Gedragwetenschappen.

Dr. Y.H.A. Bouman, senior onderzoeker Transfore.

J.C. Hutten, MSc., onderzoeker de Waag, onderdeel van De Forensische Zorgspecialisten.

Dr. F.C.A. van der Veeke, onderzoeker Kijvelanden (Fivoor Wetenschap en Behandelinnovatie)/Fivoor.

Met dank aan Antoinette Alberink, Judith Dekker, Floor van den Hanenberg, Lianne Harder, Joyce Eimers en Claudia van der Put, voor hun eerdere bijdrage aan de ontwikkeling van en/of onderzoek naar de FARE.



UNIVERSITEIT VAN AMSTERDAM

Inhoudsopgave

1. Samenvatting/Abstract	5
2. Inleiding	7
2.1 Achtergrond (heading 2/ kop 2).....	7
Interbeoordelaarsbetrouwbaarheid	7
Convergente validiteit	7
Veranderbaarheid van de FARE.....	8
Predictieve validiteit en actuariële recidiverisico-inschatting	8
Evaluatie van de bruikbaarheid van de FARE	9
2.3 Leeswijzer	9
3. Algemene onderzoeksmethode (heading 1/kop 1).....	10
3.1 Instellingen (heading 2/ kop 2)	10
3.2 Instrumenten (heading 2/ kop 2).....	10
3.3 Onderzoeksgroepen	13
3.4 Procedure.....	13
3.5 Toelichting op het onderzoek en toestemming	14
3.6 Analyses	15
4. Interbeoordelaarsbetrouwbaarheid.....	16
4.1 Methode.....	16
Onderzoeksgroepen	16
Beoordelaars	16
Interviewprotocol	16
Procedure	17
Power.....	17
Analyses	17
4.2 Resultaten.....	17
Student versus behandelaar	17
Student versus student.....	18
5. Convergente validiteit	19
5.1 Methode.....	19
Onderzoeksgroep	19
Instrumenten	19
Procedure	19
Power analyse	19
Analyses	19
5.2 Resultaten.....	20
6. Veranderbaarheid FARE	22
6.1 Methode.....	22

Onderzoeksgroep	22
Analyses	22
Missing values	23
Kanskapitalisatie.....	24
6.2 Resultaten.....	24
Reliable change index (RCI).....	31
7. Predictieve validiteit	33
7.1 Methode.....	33
Onderzoeksgroep	33
Recidivecijfers	33
Analyses	34
Kanskapitalisatie.....	34
7.2 Resultaten.....	34
Receiver Operator Curve.....	35
Cox regressieanalyses	38
8. Actuariële inschatting van het recidiverisico	48
Onderzoeksgroep	48
Analyses	48
Resultaten	48
9. Evaluatie bruikbaarheid van de FARE.....	50
9.1 Methode.....	50
9.2 Resultaten.....	50
10. FARE versie 2	53
10.2 Methode.....	53
Onderzoeksgroep	53
Beoordelaars en procedure.....	53
Power	53
Analyses	53
Resultaten	53
11. Conclusie	55
12. Referenties.....	59
Bijlage 1. Overeenkomende items FARE, IFpBE, LS/CMI en RAF GGZ	63
Bijlage 2 Overzicht wijzigingen FARE V2	66

1. Samenvatting/Abstract

In opdracht van Kwaliteit Forensische Zorg (KFZ) werd de interbeoordelaarsbetrouwbaarheid, de convergente validiteit, de veranderbaarheid, de predictieve validiteit en de praktische bruikbaarheid van de Forensische Ambulante Risico Evaluatie (FARE) onderzocht.

De data werden verzameld in de periode 1 maart 2017 tot 1 september 2019 en zijn afkomstig van volwassen cliënten in behandeling bij De Waag, Fivoor ambulante centra, Kairos, Transfore, en Verslavingszorg Noord Nederland.

Interbeoordelaarsbetrouwbaarheid. In totaal werd de FARE van 84 cliënten door studenten en behandelaren koppels beoordeeld. De Intraclass Correlatie Coëfficiënt tussen de studenten en behandelaren varieerde van *slecht* met een $ICC = ,42$ voor de risicofactor beperkte vrijetijdsbesteding tot *uitstekend* met een $ICC = ,91$ voor de risicofactor problematisch middelengebruik in het verleden. De interbeoordelaarsbetrouwbaarheid tussen studenten op de FARE versie 1 varieerde tussen slecht en uitstekend met een $ICC = ,15$ voor de dynamische risicofactor regelovertrekend gedrag in de afgelopen 6 maanden tot $ICC = ,95$ voor Aantal eerdere en huidige veroordelingen.

Convergente validiteit. Onderzoek werd uitgevoerd naar de samenhang tussen de FARE en het IFpBE, de LS/CMI en de RAF GGZ. Geen van de instrumenten gold als gouden standaard voor het ambulante forensische veld. De correlaties tussen de FARE en de andere instrumenten werden berekend om te bepalen in welke mate de instrumenten dezelfde constructen meten. De convergente validiteit is over het algemeen goed gezien de matige tot sterke samenhang tussen de FARE-items en de inhoudelijk vergelijkbare items van de andere instrumenten ($,30 \leq r \leq ,92$).

Veranderbaarheid. Op basis van de herhaalmetingen van 2180 cliënten is de veranderbaarheid van de dynamische FARE-items onderzocht. Met de FARE werden op alle dynamische risicofactoren significante, maar kleine veranderingen gemeten. De sterkste verbeteringen werden gevonden op de risicofactoren problematische (ex-)partnerrelatie, disfunctionele oplossingsvaardigheden en de gestructureerde klinische inschatting van het recidiverisico. Bij een afname van een tot drie punt(en) op een risicofactor is er sprake van een statistisch betrouwbare verbetering.

Predictieve validiteit. De predictieve validiteit van de totaalscore van de FARE-item is in de ROC-analyse de beste voorspeller van algemene recidive. Ook in de Cox regressieanalyse leverde de combinatie van alle risicofactoren tezamen een significante voorspelling van recidive op. De recidive-data werd tevens gebruikt om een actuariële risicoclassificatie op basis van alle FARE-items te formuleren.

De onderzoeksresultaten resulteerden in de FARE versie 2, met verbeterde scoringsrichtlijnen, de toevoeging van een actuariële inschatting van het algemene recidiverisico en een instructie voor het bepalen van het algemene recidiverisico.

Commissioned by Quality Forensic Care (KFZ), the interrater reliability, convergent validity, changeability, predictive validity and practical usability of the Forensic Outpatient Risk Assessment (FORE) were studied.

Data was collected between March 1st 2017 and September 1st 2019. Adults were included in the study if they were treated at the forensic outpatient treatment centers: De Waag, Fivoor, Kairos, Transfore, or Verslavingszorg Noord Nederland (Addiction Care).

Interrater reliability. In total, the FORE of 84 clients were assessed by students and therapists. The Intraclass Correlation Coefficient between students and therapists ranged from poor with an $ICC = .42$ for the risk factor Limited leisure activities to excellent with an $ICC = .91$ for the risk factor Substance abuse in the past. The interrater reliability between students on the FORE version 1 ranged between poor and excellent with an $ICC = .15$ for the dynamic risk factor Rule-breaking behavior in the past 6 months to $ICC = .95$ for the static risk factor Number of past and current convictions.

Convergent validity. The relationship between the FORE and the IFpBE, the LS/CMI and the RAF MH was analyzed. None of the instruments served as a golden standard for the outpatient forensic psychiatry. The correlations between the FORE and the other instruments were calculated to determine the extent to which the instruments measure similar constructs. The convergent validity was generally good given the moderate to strong correlation between the FORE items and the items with comparable content of the other instruments ($.30 \leq r \leq .92$).

Measuring change. Based on the repeated measures of 2180 clients, the changeability of the dynamic risk factors of the FORE were studied. The FORE measured small but significant changes in

all dynamic risk factors. The greatest improvement was measured with the risk factors Problematic (ex-)partner relationship, Dysfunctional coping and the structured professional judgement of the risk of recidivism. A decrease of one to three point(s) on a risk factor is considered a statistically reliable improvement.

Predictive validity. The predictive validity of the total score of all FORE items was the best predictor of general recidivism in the ROC analysis. In the Cox regression analysis, the combination of all risk factors also produced a significant prediction of recidivism. The recidivism data was also used to formulate an actuarial risk classification based on all FORE items.

The FORE version 2 was formulated as a result of the study outcomes: the scoring guidelines were improved, an actuarial risk classification for the assessment of risk of general recidivism was added and the instruction for determining the risk of general recidivism was extended.

2. Inleiding

2.1 Achtergrond

In de forensische ggz is het belangrijk dat informatie wordt verkregen over probleem-/risicogedrag dat gerelateerd is aan (dreigend) delictgedrag. Deze informatie wordt vaak verkregen uit risicotaxaties. Deze worden daarom naast instrumenten gericht op het meten van onder andere kwaliteit van leven en (mate van) stoornissen ook ingezet voor het monitoren van de vooruitgang in de behandeling, ook wel Routine Outcome Monitoring (ROM) genoemd. Sinds enkele jaren wordt de Forensisch Ambulante Risico Evaluatie (FARE) gebruikt in het ambulante forensische veld om inzicht te krijgen in (veranderingen over tijd in) het klachtenpatroon. Dat gebeurt met de 11 dynamische risicofactoren van de FARE. In de FARE zijn ook zes statische risicofactoren opgenomen om het volgende in kaart te brengen: de delictgeschiedenis, of er sprake is van een langdurig patroon van problemen op het gebied van opleiding/werk en middelengebruik, en om het recidiverisico in te kunnen schatten op basis van het heden en verleden. De FARE beoogt daarmee een ROM-instrument en een risicotaxatie-instrument te zijn.

In de voorliggende studie worden de volgende deelvragen over de psychometrische kwaliteiten van FARE als ROM-, en risicotaxatie-instrument beantwoord:

- Wat is de overeenstemming tussen beoordelaars op de FARE (interbeoordelaarsbetrouwbaarheid)?
- Wat is de samenhang van de FARE met andere risicotaxatie-instrumenten die hetzelfde beogen te meten (convergente validiteit)?
- In welke mate wordt in herhaalmetingen met de FARE verandering bij cliënten gemeten?
- In hoeverre zijn de risicofactoren van de FARE afzonderlijk en de FARE als geheel voorspellend voor recidive?
- Zijn er subgroepen en/of cliëntprofielen te onderscheiden en zijn er verschillen in predictieve validiteit tussen subgroepen?
- Is er aanleiding om op basis van de data verzameld met de FARE de actuariële risico inschatting bij te stellen?

Voordat we deze deelvragen beantwoorden, wordt eerst de wetenschappelijke stand van zaken van de in Nederland veelgebruikte risicotaxatie-instrumenten nader besproken.

Interbeoordelaarsbetrouwbaarheid

Voor een betrouwbaar risicotaxatie-instrument en ROM-instrument geldt dat de uitkomst onafhankelijk moet zijn van de beoordelaar. Wanneer twee beoordelaars een cliënt beoordelen met hetzelfde instrument op basis van dezelfde informatie, behoren ze tot overeenkomstige scores te komen (Harte & Breukink, 2010). De mate van overeenstemming kan worden uitgedrukt in een Intraclass Correlatie Coëfficiënt (ICC) die kan variëren tussen de 0 (nul) en 1.

De interbeoordelaarsbetrouwbaarheid van de risicotaxatie-instrumenten die in Nederland het meest worden gebruikt, is over het algemeen voldoende tot zeer goed (Blok, De Beurs, De Ranitz, & Rinne, 2010; Van Horn, Eisenberg, & Uzieblo, 2016). Bijvoorbeeld voor de Historical Clinical Risk Management 20 (HCR-20) varieerde de ICC tussen matig ($ICC = ,55$) en zeer goed ($ICC = ,89$) (Hildebrand, Hesper, Spreen, & Nijman, 2005; Vogel, Ruiters, Hildebrand, Bos, & Ven, 2004; Vogel & Ruiters, 2005). Voor de Historische Klinische Toekomstige 30 (HKT-30) varieerde de ICC tussen matig en goed ($,61 \leq ICC \leq ,87$) (Canton, Van der Veer, Van Panhuis, Verheul, & Van der Brink, 2004; Hildebrand et al., 2005; Schönberger et al., 2008) en voor de Historische Klinische Toekomstige Revised versie (HKT-R), de opvolger van de HKT-30, lag de ICC voor de afzonderlijke schalen en het totale instrument tussen matig ($ICC = ,42$) en zeer goed ($ICC = ,85$) (Bogaerts, Spreen, Ter Horst, & Gerlsma, 2018). Voor de Level of Service Case Management Inventory (LS/CMI) lag de ICC tussen goed en zeer goed ($,65 \leq ICC \leq ,95$) (Andrews & Bonta, 1995; Rocque & Plubber-Beale, 2014). Wat betreft de interbeoordelaarsbetrouwbaarheid van het Instrument voor Forensische Behandeling Evaluatie (IFBE) is de ICC van de afzonderlijke items en drie schalen substantieel ($ICC > ,65$, Schuringa, Spreen, & Bogaerts, 2014). Tot slot toont onderzoek naar het Risicotaxatietaxatie-instrument voor de Ambulante Forensische Geestelijke Gezondheidszorg (RAF GGZ) jeugd aan dat de interbeoordelaarsbetrouwbaarheid zeer goed is ($ICC = ,78$, Van Horn, Wilpert, Bos, & Mulder, 2009).

Convergente validiteit

De convergente validiteit is onderdeel van de constructvaliditeit die aangeeft of en in welke mate een instrument inhoudelijk overeenkomt met de constructen die het beoogt te meten (Van

der Knaap, Leenarts, & Nijssen, 2007). De convergente validiteit geeft de mate van overeenstemming weer tussen twee instrumenten die, in theorie, hetzelfde construct beogen te meten. Om de convergente validiteit te bepalen, werd de FARE vergeleken met het Instrument voor Forensisch poliklinische Behandel Evaluatie (IFpBE), de LS/CMI en de RAF GGZ.

De convergente validiteit wordt meestal uitgedrukt in correlaties. De volgende richtlijnen worden vaak gehanteerd: een correlatie tot $r < ,30$ is zwak, een correlatie van $,30 \geq r \leq ,49$ is middelsterk en vanaf $r \geq ,50$ spreken we van een sterke samenhang (Van der Knaap, Leenarts, & Nijssen, 2007).

Omdat de FARE gebaseerd is op onderzoek naar de correlatie tussen de items van het IFpBE, de LS/CMI en de RAF GGZ en alle vier de instrumenten recidiverisico/behandelvoortgang en onderliggende risicofactoren meten, wordt een sterke samenhang ($r \geq ,50$ of hoger) verwacht tussen de items en recidiverisico-inschattingen van de instrumenten (Van Horn et al., 2016a).

Veranderbaarheid van de FARE

Zoals eerder aangegeven is de FARE ontwikkeld voor Routine Outcome Monitoring (ROM) én om het recidiverisico in te schatten. Ten behoeve van de ROM wordt de behandelvoortgang bepaald op basis van veranderingen op de dynamische risicofactoren (D1 t/m D11). Een voorbeeld van een onderzoek naar ROM in de forensische ggz is de studie van Nijman, De Kruyk en Van Nieuwenhuizen (2004), waarin ROM-metingen van de HKT-30 in forensisch psychiatrische centra werden onderzocht. Ze vonden dat de meeste gedragsveranderingen plaatsvonden in de eerste 14 maanden van behandeling. Van der Veeken, Lucieer en Bogaerts (2018) onderzochten de behandelvoortgang in twee forensisch psychiatrische centra met het IFBE (klinische variant van het IFpBE). Zij vonden meer behandelvoortgang bij cliënten die aan de start van de behandeling veel probleemgedrag, weinig beschermend gedrag en resocialisatie vaardigheden vertoonden dan bij cliënten die weinig probleemgedrag en veel beschermend gedrag en resocialisatie vaardigheden vertoonden.

In het voorliggende onderzoek wordt verwacht dat de FARE sensitief genoeg is waardoor relatief kleine veranderingen zichtbaar worden.

Predictieve validiteit en actuariële recidiverisico-inschatting

De FARE als ROM en Risicotaxatie-instrument sluit aan bij het leidende theoretisch kader voor een effectieve rehabilitatie, het Risk-Need-Responsivity model (RNR, Bonta & Andrews, 2017). De FARE levert een inschatting op van het recidiverisico om de intensiteit en duur van de behandeling (Risk) op af te stemmen en op basis van de dynamische risicofactoren kunnen behandeldoelen (Needs) worden geformuleerd. De *Central Eight* (C8) criminogene factoren worden verondersteld het sterkt met recidive samen te hangen voor alle forensische (sub-)populaties (Bonta, Blais, & Wilson, 2014; Eisenberg et al., 2019; Grieger & Hossner, 2014; Gutierrez, Wilson, Rugge, & Bonta, 2013; Olver, Stockdale, & Wormtith, 2014; Wooditch, Tang, & Taxman, 2014). Waarbij wel rekening moet worden gehouden met individuele verschillen. Zo werd gevonden dat problematisch middelen gebruik en persoonlijke/emotionele problemen recidive voor vrouwen sterker voorspelden dan voor mannen (Olver et al., 2014) en beperkte vrijetijdsbesteding voorspelde recidive sterker voor personen die behoorden tot de etnische minderheid dan personen die behoorden tot de etnische meerderheid (Gutierrez et al., 2013; Olver et al., 2014; Wilson & Gutierrez, 2014). Veranderingen in deze C8 risicofactoren zouden dus moeten leiden tot veranderingen in het recidiverisico en daarmee tot veranderingen in feitelijke recidive. In welke mate een risicotaxatie-instrument toekomstig recidive kan voorspellen, wordt onderzocht met de predictieve validiteit.

Tot nu toe is in Nederland in beperkte mate onderzoek gedaan naar de predictieve validiteit of voorspellende waarde van risicotaxatie-instrumenten specifiek voor de ambulante forensische populatie. Bouman, Schene en De Ruiters (2008) vonden binnen een ambulante forensische populatie een beschermend effect van subjectief welbevinden op zelfgerapporteerde recidive. En een significante correlatie tussen een hoog recidiverisico op de LSI-R en zelf gerapporteerde recidive (binnen drie maanden). Door Troquete en collega's (2015) werd de predictieve validiteit op korte termijn (binnen drie en zes maanden) van de Short-Term Assessment of Risk and Treatability (START) onderzocht bij een ambulante forensische populatie. De START voorspelde recidive op basis van het gestructureerd klinisch oordeel binnen drie en zes maanden matig, met respectievelijk een *Area Under the Curve* (AUC) van ,64 en ,67. Ten behoeve van de ontwikkeling van de FARE is de voorspellende waarde van de RAF GGZ onderzocht. De voorspellende waarde van de gestructureerde klinische inschatting met de RAF GGZ was matig met een AUC = ,60.

Daarnaast werd de voorspellende waarde van de somscore, de actuariële inschatting, onderzocht en deze leverde een betere voorspelling op met een AUC van ,71 (Van Horn, et al., 2016).

Zoals besproken door Guy, Douglas en Hart (2015) wordt al lang gediscussieerd door ontwikkelaars van risicotaxatie-instrumenten en onderzoekers wat de beste benadering is voor het taxeren van het recidiverisico. Zo was Meehl in 1954 de eerste die beargumenteerde dat een statistische benadering (actuariële inschatting) de voorkeur verdient boven de klinische inschatting. Ægisdóttir, Spengler en White (2006) vonden in een meta-analyse van studies naar de predictieve validiteit van de klinische versus actuariële inschatting dat een actuariële inschatting een betere voorspelling opleverde. Deze onderzoeken waren gericht op de ongestructureerde klinische inschatting en niet op de gestructureerde klinische inschatting zoals tegenwoordig gebruikelijk is in de uitvoering van risicotaxatie. In de LS/CMI (Bouman, Wismeyer, Lammers, & Philipse, 2012) wordt voor de inschatting van het recidiverisico gebruik gemaakt van een actuariële benadering met aanvullend de mogelijkheid voor een '*clinical override*' waarbij de behandelaar ervoor kan kiezen om de actuariële inschatting bij te stellen. Uit een Canadese studie naar de predictieve validiteit van de LS/CMI waarbij een onderscheid werd gemaakt tussen de actuariële benadering en de '*clinical override*' bleek dat beoordelaars bij het gebruik van de '*clinical override*' de actuariële inschatting van het recidiverisico meestal naar boven bijstelde en de eindconclusie was dat de actuariële benadering een betere predictieve validiteit opleverde (Wormith, Hogg, & Guzzo, 2012).

In deze deelstudie werd onderzocht of de individuele risicofactoren van de FARE en de FARE als geheel voorspellend zijn voor algemene recidive. De FARE bestaat grotendeels uit de C8 risicofactoren. Daarom is de verwachting dat de voorspellende waarde van de FARE op of boven het niveau ligt van andere risicotaxatie-instrumenten die deze C8 in meer of mindere mate meten. Daarnaast is de predictieve validiteit onderzocht voor subgroepen: mannen, vrouwen, jongvolwassenen, cliënten die geboren zijn in en buiten Nederland, cliënten in behandeling voor agressie en/of vermogensdelicten, huiselijk geweld of seksueel grensoverschrijdend gedrag. De predictieve validiteit van risicotaxatie-instrumenten bij subgroepen op basis van genoemde kenmerken is nog niet eerder gedaan. De verwachting is dat de subgroepen over het algemeen weinig variatie zullen laten zien in de mate waarin de FARE, recidive accuraat voorspelt.

In de FARE is een actuariële inschatting van het recidiverisico opgenomen op basis van de statische risicofactoren. Dat was mogelijk omdat we uit andere risicotaxatie-instrumenten (RAF en Risc) een grote dataset tot onze beschikking hadden die vergelijkbare statische risicofactoren meten (zie Van Horn, 2016a voor een uitgebreide toelichting). Met de huidige dataset van de FARE is het ook mogelijk om een actuariële classificatie te berekenen van de dynamische risicofactoren van de FARE en van alle FARE-risicofactoren (statisch en dynamisch). Hierover wordt ook gerapporteerd in dit rapport.

Evaluatie van de bruikbaarheid van de FARE

Bij de ontwikkeling van de FARE werd door gebruikers aangegeven dat het belangrijk was dat de FARE een kort en praktisch bruikbaar instrument zou worden (Van Horn et al., 2016b). In dit rapport is een beschrijving opgenomen van de evaluatie naar de praktische bruikbaarheid van de FARE.

2.3 Leeswijzer

In hoofdstuk 3 worden de instrumenten en procedure beschreven die voor de beantwoording van meerdere deelvragen van toepassing zijn. Beschreven worden de betrokken forensische poliklinieken, de gebruikte instrumenten, informatieverstrekking aan betrokken cliënten en het verlenen van toestemming door cliënten.

De onderzoeksmethoden en resultaten worden verder per deelvraag uitgewerkt in respectievelijk de hoofdstukken 4 'Interbeoordelaarsbetrouwbaarheid', 5 'Convergente validiteit', 6 'Veranderbaarheid' en 7 'Predictieve validiteit'. In hoofdstuk 8 beschrijven we de aanvullende actuariële risicoclassificatie van het recidiverisico en in hoofdstuk 9 de evaluatie van behandelaren van de FARE versie 1. De resultaten van de deelvragen (psychometrische kwaliteit en praktische bruikbaarheid) hebben geresulteerd in een nieuwe versie van de FARE. In hoofdstuk 10 wordt de FARE versie 2 beschreven. Tot slot wordt in hoofdstuk 11 de eindconclusie gegeven.

3. Algemene onderzoeksmethode

In dit hoofdstuk wordt de onderzoeksmethode beschreven voor alle deelvragen, zoals de participerende instellingen, de gebruikte instrumenten, en de procedures

3.1 Instellingen

De Waag is sinds 1992 een ambulante forensische zorgaanbieder en is onderdeel van De Forensische zorgspecialisten. De Waag biedt zowel behandeling op vrijwillige basis als behandeling binnen een strafrechtelijk kader. De Waag heeft acht vestigingen en daarnaast twee satellietlocaties. Deze liggen verspreid over de Randstad en Midden-Nederland. Bij de Waag werken ongeveer 300 mensen, voornamelijk gz-psychologen, psychotherapeuten, klinisch psychologen en psychiaters. De Waag heeft een jaarlijkse instroom van ongeveer 4000 cliënten.

Fivoor Ambulant (voorheen Het Dok) is een forensisch psychiatrische poli- en dagkliniek, biedt behandeling bij grensoverschrijdend gedrag en is onderdeel van Fivoor. Fivoor Ambulant biedt zowel behandeling op vrijwillige basis als behandeling in een strafrechtelijk kader. Fivoor Ambulant bestaat uit 8 ambulante centra en is een tweedelijns zorginstelling. Binnen de vestigingen van Tilburg, Breda, Dordrecht en Rotterdam is data verzameld voor dit onderzoek. Deze vier vestigingen van Fivoor Ambulant hebben een jaarlijkse instroom van ongeveer 1400 cliënten. Er werken ongeveer 80 behandelaren, zoals psychiaters, psychologen en verpleegkundigen.

Kairos is een forensisch psychiatrische polikliniek met vestigingen in vijf locaties in Zuid-Gelderland en Oost Noord-Brabant. Kairos biedt zowel behandeling op vrijwillige basis als behandeling in strafrechtelijk kader. Ongeveer 90 behandelaren zien jaarlijks 1200 patiënten. Kairos is onderdeel van de Pompestichting en bestaat sinds 1992.

Transfore (onderdeel van de Dimence Groep) biedt sinds 1994 poliklinische behandeling aan forensische psychiatrische patiënten onder de naam 'de Tender'. Behandeling bij de Tender vindt zowel in een strafrechtelijk kader als op vrijwillig basis plaats. De Tender heeft vestigingen in Overijssel en Gelderland. De Tender behandelt jaarlijks ongeveer 1000 cliënten en daarnaast zijn 180 cliënten in behandeling bij ForFACT teams. In totaal werken 75 behandelaren bij de zes vestigingen van De Tender en de ForFACT teams van Transfore.

Verslavingszorg Noord Nederland (VNN) heeft sinds 2011 een forensische polikliniek waar consult, diagnostiek en behandeling wordt geboden aan mensen met een verslaving die een strafbaar feit hebben gepleegd of hiervan worden verdacht. Cliënten bij VNN hebben een (gedeeltelijk) voorwaardelijke straf om mee te werken aan psychiatrisch onderzoek/behandeling. VNN heeft vestigingen in de provincies Groningen, Friesland en Drenthe.

3.2 Instrumenten

De Forensische Ambulante Risico Evaluatie (FARE; Van Horn, Eisenberg, Bouman, Van den Hanenberg, & Bogaerts, 2016c) is een risicotaxatie- en behandel-evaluatie instrument dat is ontwikkeld voor gebruik in de ambulante forensische psychiatrische zorg. De FARE is een generiek instrument en kan worden gebruikt bij alle cliënten van 18 jaar en ouder, ongeacht het type indexdelict. De FARE bestaat uit 17 items die worden gescoord op een 5-puntsschaal (0-4), zes statische risicofactoren en 11 dynamische risicofactoren (zie Tabel 1).

De actuariële risico-inschatting van de FARE volgt uit de somscore van de zes statische items (somscore 0 – 24). Deze somscore correspondeert met één van de vijf risiconiveaus (zeer laag, laag, matig, hoog, zeer hoog). Vervolgens geeft de behandelaar een (gestructureerde) klinische inschatting van het recidiverisico, waarbij zowel de statische als de dynamische items worden meegewogen. De actuariële scoring op basis van de statische risicofactoren is gebaseerd op secundaire analyses van onderzoeksgegevens naar de Risicotaxatie-instrument voor de Ambulante Forensische GGZ (RAF GGZ). Een uitgebreide beschrijving van de ontwikkeling van de actuariële scoring en de ontwikkeling van de FARE wordt gegeven in het KFZ-rapport 'Kernset K-factoren voor het ambulante forensische veld' (Van Horn et al., 2016a).

Tabel 1. Items van de FARE

Statische risicofactoren	Dynamische risicofactoren
S1 Leeftijd eerste politiecontact	D1 Disfunctioneren opleiding/werk
S2 Aantal eerdere en huidige veroordelingen	D2 Financieel wanbeleid,
S3 Aantal type delicten	D3 Delinquent sociaal netwerk
S4 Regelovertredend gedrag in het verleden	D4 Beperkte vrijetijdsbesteding
S5 Instabiliteit opleiding/werk in het verleden	D5 Problematische (ex-) partnerrelatie
S6 Problematisch middelengebruik in het verleden	D6 Instabiliteit woonsituatie
	D7 Problematisch middelengebruik
	D8 Gebrekkige impulsbeheersing
	D9 Disfunctionele oplossingsvaardigheden
	D10 Antisociale houding
	D11 Regelovertredend gedrag

Het Instrument voor Forensische poliklinische Behandel Evaluatie (IFpBE) is afgeleid van het IFBE. Dit instrument werd oorspronkelijk ontwikkeld voor klinische doeleinden en vervolgens vertaald naar een poliklinische setting (Schuringa, 2011; Schuringa, Spreen, & Bogaerts, 2014). Het IFpBE bestaat uit 22 items (zie Tabel 2) waarvan 14 items overeenkomen met de klinische items van de HKT-R (Spreen, Brand, Ter Horst, & Bogaerts, 2014; Bogaerts et al., 2018). Het instrument kan worden gescoord op een 17-puntsschaal met vijf ankerpunten. Voor onderzoeksdoeleinden kunnen de scores op alle items worden opgeteld (range 17 – 374). Behandelvoortgang wordt gemonitord door op itemniveau verandering in scores van meerdere metingen te vergelijken.

Het IFpBE is nog niet onderzocht op psychometrische kwaliteiten, maar uit studies naar het IFBE blijkt een goede interne consistentie voor de drie factoren (probleemgedrag, $\alpha = ,86$, beschermend gedrag, $\alpha = ,90$ en resocialisatie vaardigheden, $\alpha = ,88$). Test-herstest betrouwbaarheid van de drie factoren was erg goed (problematisch gedrag, $\alpha = ,85$, beschermend gedrag, $\alpha = ,87$ en resocialisatie vaardigheden, $\alpha = ,89$) (Schuringa et al., 2014). Tevens heeft het IFBE een goede predictieve validiteit van de drie IFBE factoren voor de voorspelling van algemeen geweld in de kliniek binnen een groep patiënten met een persoonlijkheidsstoornis, psychotische stoornis of aan middelen gebonden stoornis (beschermend gedrag, $AUC = ,69$, probleemgedrag, $AUC = ,77$ en resocialisatiegedrag, $AUC = ,75$) en fysiek geweld (beschermend gedrag, $AUC = ,65$, probleemgedrag, $AUC = ,81$ en resocialisatiegedrag, $AUC = ,76$) (Van der Veeken, Lucieer, & Bogaerts, 2016). In een andere studie blijken twee van de drie factoren een gemiddelde tot goede voorspelling te hebben van druggebruik in de kliniek (beschermend gedrag, $d = ,57$, $p < ,05$; probleemgedrag, $d = 1,47$, $p < ,001$) (Schuringa, Heininga, Spreen, & Bogaerts, 2018).

Tabel 2. Items van het IFpBE (corresponderende K-items van de HKT-R)

Items 1 – 11	Items 12 - 22
1. Probleeminzicht (HKT-R item K1)	12. Antisociaal gedrag (K5)
2. Meewerken aan behandeling (K9)	13. Vijandigheid (K6)
3. Verantwoordelijkheid voor het delict (K10)	14. Seksueel grensoverschrijdend gedrag
4. Copingvaardigheden (K11)	15. Manipulatief gedrag
5. Evenwichtige dagindeling	16. Schending van voorwaarden (K12)
6. Arbeidsvaardigheden (K13)	17. Antisociaal netwerk (K14)
7. Sociale vaardigheden (K7)	18. Medicatietrouw
8. Zelfredzaamheid (K8)	19. Psychotische symptomen (K2)
9. Financiële vaardigheden	20. Partnerrelatie
10. Administratieve vaardigheden	21. Middelen (K3)
11. Impulsiviteit (K4)	22. Fysiek agressief gedrag

De Level of Service Case Management Inventory (LS/CMI; Bouman, Wismeyer, Lammers, & Philipse, 2012) is de meest recente versie van de Level of Service serie. De voorloper van de LS/CMI, de Level of Service Inventory-revised (LSI-r; Andrews & Bonta, 1995) is een instrument voor het inventariseren van de risico- en behoeftefactoren van het Risk Need Responsivity model

(RNR-model; Bonta & Andrews, 2017) bij volwassen delinquenten. Het bestaat uit 54 items verdeeld over de volgende 10 domeinen (zie Tabel 3).

Tabel 3. Items van de LS/CMI 54 items verdeeld over 10 domeinen

Domeinen 1 – 6	Domeinen 7 - 10
<p>1. Delictgeschiedenis</p> <p>1. Eerdere veroordeling(en) gehad?</p> <p>2. Twee of meer eerdere veroordelingen gehad?</p> <p>3. Drie of meer eerdere veroordelingen gehad</p> <p>4. Drie of meer huidige delicten?</p> <p>5. Arrestaties voor 16 jaar?</p> <p>6. Ooit gevangen gezeten na veroordeling?</p> <p>7. Ooit bestraft voor wangedrag in een inrichting?</p> <p>8. Aanklacht of schending van reclasseringsvoorwaarden?</p> <p>2. Opleiding/Werk <i>Indien op de arbeidsmarkt:</i></p> <p>9. Op dit moment minder dan 20 uur betaald werk per week?</p> <p>10. Vaak werkloos?</p> <p>11. Nooit een vol jaar aan het werk geweest?</p> <p><i>Schooltijd of indien op school:</i></p> <p>12. 3 jaar of minder voorgezet onderwijs?</p> <p>13. 5 jaar of minder voorgezet onderwijs?</p> <p>14. Ooit geschorst of van school gestuurd?</p> <p>15. Deelname/prestatie</p> <p>16. Omgang met medeleerlingen/collegae</p> <p>17. Omgang met leraren/superieuren</p> <p>4. Familie/Relatie</p> <p>18. Mate van tevredenheid over huidige leefvorm</p> <p>19. Relatie met ouders</p> <p>20. Relatie met andere familieleden</p> <p>21. Strafblad in de naaste familie?</p> <p>5. Vrije tijd/ontspanning</p> <p>22. Geen recente deelname aan clubs/organisaties</p> <p>23. Mate van actieve vrijetijdsbesteding</p> <p>6. Vrienden/Kennissen</p> <p>24. Enige bekenden met crimineel gedrag</p> <p>25. Enige vrienden met crimineel gedrag</p> <p>26. Heeft minder dan drie bekenden zonder crimineel gedrag</p> <p>27. Heeft minder dan drie vrienden zonder crimineel gedrag</p>	<p>7. Alcohol/Drugsproblemen</p> <p>28. Ooit een drankprobleem gehad</p> <p>29. Ooit een drugsprobleem gehad</p> <p>30. Mate van actuele drankproblemen</p> <p>31. Mate van actuele drugsproblemen.</p> <p>32. Relatie drank/drugs met criminaliteit?</p> <p>33. Familie klaagt over drank/drugsgebruik</p> <p>34. Problemen op school/werk door drank/drugsgebruik.</p> <p>35. Medische problemen als gevolg van drank/drugsgebruik of andere indicatoren voor drank/drugsgebruik</p> <p>8. Procriminele attitude</p> <p>36. Goedkeuring van delict(en)</p> <p>37. Afwijzing van maatschappelijke conventies</p> <p>38. Weinig begrip voor vonnis</p> <p>36. Goedkeuring van delict(en)</p> <p>39. Weinig begrip voor toezicht/behandelmaatregel</p> <p>9. Antisociaal Patroon</p> <p>40. (Vermoeden van) Antisociale trekken</p> <p>41. Vroegtijdig en gevarieerd antisociaal gedrag</p> <p>a) Ernstige aanpassingsproblemen in de kindertijd, zoals aangegeven door school en/of maatschappelijke instanties, of gearresteerd/veroordeeld voor 16e jaar</p> <p>b) Officiële aangifte van (dreiging met) geweld</p> <p>c) Ooit ontsnapt of geprobeerd te ontsnappen uit een instelling</p> <p>d) Schenden voorwaarde reclasseringstoezicht</p> <p>42. Criminele attitude</p> <p>a) Goedkeuring delict(en)</p> <p>b) Afwijzing van maatschappelijke conventies</p> <p>c) Weinig begrip voor toezicht/behandelmaatregel</p> <p>43. Problemen op meerdere levensgebieden</p> <p>a) Financiële problemen</p> <p>b) Minstens 3 adreswijzigingen het afgelopen jaar.</p> <p>c) Nooit een vol jaar aan het werk geweest?</p> <p>d) 3 jaar of minder voortgezet onderwijs?</p> <p>e) Ooit geschorst of van school gestuurd?</p> <p>f) Onbevredigende relatie met ouders</p> <p>g) Onbevredigende mate van actieve vrijetijdsbesteding</p> <p>h) Heeft minder dan drie vrienden zonder crimineel gedrag</p>

De items worden dichotoom (ja/nee) gescoord of op een 4-puntsschaal variërend van 0=zeer onbevredigend tot 3=zeer bevredigend. De items worden per domein opgeteld, waarbij een ja-score en een 0 of 1 score als 1 punt wordt geteld. Vervolgens worden alle scores opgeteld met een range van 0 – 54. De LSI-r wordt gescoord op basis van dossiers en een interview met de cliënt.

De gemiddelde predictieve validiteit van de LSI-r is $r = ,36$ voor algemene recidive ($k > 33$) en $r = ,25$ voor gewelddadige recidive ($k > 16$; Andrews, Bonta, & Wormith, 2006). Campbell, French en Gendreau (2007) rapporteerden $Z+ = ,28$ (CI = $,25 - ,31$; $k = 19$) voor gewelddadige recidive. Inzake de interbeoordelaarsbetrouwbaarheid, zoals gerapporteerd door de auteurs zelf (Andrews & Bonta, 1995), varieerde de ICC van $,80 - ,95$. Ander onderzoek laat een lagere ICC zien (Rocque & Plummer-Beale, 2014; ICC-totaalscore = $,65$).

In de LS/CMI zijn de 54 LSI-r items gecombineerd en geherformuleerd in de 43 items van LS/CMI sectie 1, waarbij de inhoud van de items hetzelfde is gebleven. Om deze reden werd in dit

onderzoek enkel sectie 1 onderzocht. De psychometrische eigenschappen van de LS/CMI zijn nog niet eerder onderzocht, maar de verwachting is dat deze vergelijkbaar zijn met de LSI-R.

Het Risicotaxatie-instrument voor de Ambulante Forensische GGZ (RAF GGZ volwassenen; Van Horn, Wilpert, Scholing, & Mulder, 2008) is een gestructureerd klinisch risicotaxatie-instrument waarmee op gestandaardiseerde wijze het recidiverisico wordt ingeschat bij volwassen delinquenten vanaf 18 jaar die zijn geïndiceerd voor ambulante forensische behandeling. Het is een generiek instrument dat kan worden afgenomen ongeacht het type indexdelict. Het instrument bestaat uit 79 risicofactoren verdeeld over 12 domeinen (zie Tabel 4). De risicofactoren worden dichotoom (ja/nee) gescoord of op een 3-puntsschaal (0=niet aanwezig, 1=enigszins aanwezig, 2=aanwezig). De scoringsprocedure van de RAF GGZ verloopt in 4 stappen. Na het scoren van de items (stap 1), wordt per domein op een 6-puntschaal (0 – 5) een beoordeling gegeven van het algemeen functioneren op dat domein (stap 2). In de derde stap wordt beoordeeld of de domeinen waarop de cliënt problematisch functioneert (score 3 – 5) een relatie hebben met het delictgedrag. Als er sprake is van een relatie met het delict gedrag wordt het domein een risicofactor genoemd. In de vierde stap wordt op een 5-punt schaal (laag, laag-matig, matig, matig-hoog, hoog) een gestructureerde klinische inschatting gegeven voor alle type delicten die de cliënt in zijn criminele voorgeschiedenis heeft gepleegd. Voor onderzoeksdoeleinden kunnen de scores van de 79 items worden opgeteld (range 0 - 123). De predictieve validiteit van de RAF GGZ volwassenen bleek goed met een $AUC = ,71$ (Eisenberg, Van Horn, Van der Put, Hendriks & Stams, 2015).

3.3 Onderzoeksgroepen

Per onderzochte psychometrische eigenschap van de FARE zijn de onderzoeksgroepen anders samengesteld. Deze worden in het daartoe behorende hoofdstuk afzonderlijk besproken. De onderzoeksgegevens van de verschillende instellingen zijn geanonimiseerd samengevoegd en binnen de Waag geanalyseerd. Van de cliënten van Fivoor, Kairos en/of Transfore ontbreken gegevens zoals geslacht, leeftijd en primaire diagnose. Binnen deze instellingen wordt de Algemene Verordening Gegevensbescherming (AVG) strikter uitgevoerd dan de definitie voor het anonimiseren van onderzoeksgegevens. Genoemde instellingen gegevens zoals geslacht, leeftijd en primaire diagnose onder de informatie die herleidbaar is tot individuen (zoals naam en geboorteplaats). Deze informatie wordt daarom niet gedeeld met derden.

3.4 Procedure

Het onderzoeksvoorstel werd op 26 juni 2017 goedgekeurd door the Ethics Review Board (ERB) of the School of Social and Behavioral Sciences van Tilburg University (EC-2017.36). Onderzoekers van de Waag, Transfore en Fivoor hebben de behandelaars getraind die aan het onderzoek hebben deelgenomen. Daarnaast werden de behandelaars geïnstrueerd over de procedure voor het informeren van cliënten en het verkrijgen van schriftelijk toestemming voor deelname aan één of meerdere deelonderzoeken.

Voor een uitgebreide beschrijving van de procedure wordt verwezen naar het onderzoeksvoorstel dat door de ERB werd goedgekeurd (EC-2017.36). In de volgende paragrafen wordt de procedure op hoofdlijn toegelicht.

Tabel 4. Domeinen en items van de RAF GGZ

D1. Eerdere en huidige delicten	D8. vervolg
1a. Veroordelingen	8e. Relatie middelengebruik met criminaliteit/wangedrag
1b. Officiële aangiften	8f. Interpersoonlijke problemen met naasten door middelengebruik
1c. Niet officieel geregistreerde delicten	8g. Problemen op school/werk door middelengebruik
1d. Gebruik van wapens en/of serieuze doodsbedreigingen	8h. Medische problemen door middelengebruik
1e. Escalatie in frequentie of ernst delicten	D9. Persoonlijk / emotioneel functioneren
1f. Jonge leeftijd eerste uiting antisociaal gedrag	9a. Slachtoffer van kindermishandeling
1g. Ooit gevangen gezeten na veroordeling	9b. Als kind gepest op school
1h. Schending van voorwaarden	9c. Suïcidale / homicidale gedachten
D2. Opleiding / werk	9d. Gebrek aan zelfinzicht
2a. Opleiding voortgezet onderwijs niet afgerond	9e. Impulsiviteit
2b. Ooit geschorst of van school gestuurd	9f. Hoog niveau van ervaren stress
2c. Geringe Schoolprestaties	9g. Slechte copingvaardigheden
2d. Weinig interesse in/binding met school	9h. Problemen met hanteren van boosheid
2e. Gedragsproblemen op school	9i. Psychische problematiek (As I)
2f. Problemen met betrekking tot arbeidsverleden	9j. Persoonlijkheidsstoornissen/ -trekken (As II)
2g. Werkloos afgelopen jaar	9k. (Aanwijzingen voor) Cognitieve/ intellectuele beperkingen
2i. Werk prestaties	9l. Inadequaat gebruik medicatie
2j. Omgang met collegae	D10. Houding delictgedrag
2k. Omgang met Leidinggevende	10a. Goedkeuring/vergoelijking van delicten in het algemeen
D3. Financieel beheer	10b. Minimalisering/ontkenning van eigen delicten
3a. Ontvangt uitkering	10c. Veroordeeld voor indexdelict
3b. Schulden	D11. Hulpverlening
D4. Wonen	11a. Hulpverlening in het verleden gehad?
4a. Instabiliteit woonsituatie	11b. Blootstelling aan destabiliserende invloeden
4b. Criminaliteit woonomgeving	11c. Ontbreken van realistische toekomstplannen
D5. Intieme relatie / familie	11d. Negatieve houding ten opzichte van interventie
5a. Instabiliteit van relaties	11e. Reageert niet op behandeling
5b. Slechte Relaties met ouders / verzorgers	11f. Werkt niet mee aan behandelmogelijkheden
5c. Slechte Relatie met andere (schoon)familieleden	11g. Aanwezigheid behandeling
5d. Kinderen	11h. Inzet tijdens behandeling
5e. Politiecontact naaste familie/gezin	11i. Behandeling afgebroken
D6. Sociaal netwerk	D12. Seksuele problematiek
6a. Sociaal geïsoleerd	12a. Hands-on zedendelicten
6b. Criminele vrienden	12b. Hands-off zedendelicten
6c. Geen prosociale vrienden	12c. Seksuele deviatie
6d. Geringe beschikbaarheid persoonlijke steun	12d. Hoge dichtheid seksuele delicten
6e. Geen prosociale steun	12e. Mate van planning van seksuele delicten
D7. Vrijtijdsbesteding	12f. Meerdere type seksuele delicten
7a. Geen individuele vrijtijdsbesteding	12g. Seks als copingstrategie
7b. Geen contextuele vrijtijdsbesteding	12h. Seksuele preoccupatie
D8. Middelen problematiek	12i. Niet-verwante slachtoffers van seksuele delicten
8a. Drankmisbruik/afhankelijkheid in het verleden	12j. Onbekende slachtoffers
8b. Drugsmisbruik/afhankelijkheid in het verleden	12k. Mannelijke slachtoffers van seksuele delicten
8c. Drankmisbruik/afhankelijkheid afgelopen jaar	12l. Minderjarige slachtoffers van seksuele delicten
8d. Drugsmisbruik/afhankelijkheid afgelopen jaar	12m. Lichamelijk letsel bij slachtoffers van seksuele delicten
	12n. Aantal slachtoffers van seksuele delicten

3.5 Toelichting op het onderzoek en toestemming

Alvorens met het onderzoek te starten, gaf de behandelaar de cliënt uitleg over het onderzoek en overhandigde een informatiebrief over het onderzoek. Daarna werd de cliënt om schriftelijke toestemming gevraagd voor vrijwillige deelname aan het onderzoek. Cliënten kregen na het overhandigen van de informatiebrief tot de volgende afspraak (1 à 2 weken) met hun behandelaar de tijd om te beslissen over deelname aan het onderzoek. Na instemming werd voor het afnemen van het interviewprotocol de toestemmingsverklaring ondertekend. De behandelaar nam het

interview af en zorgde voor de audio-opname ervan. Het interview vond plaats gedurende 1-3 behandelingsessies. Na afloop van het interview kreeg de cliënt een cadeaubon ter waarde van €15,- (kort interview, zie voor toelichting paragraaf 4.1) of €25,- (lang interview, zie paragraaf 4.1). Het interviewprotocol ten behoeve van de interbeoordelaarsbetrouwbaarheid en convergente validiteit kon zowel in de beginfase van de behandeling als in een latere fase worden afgenomen (bijvoorbeeld t.b.v. de tweede ROM-meting of de eind ROM-meting). Cliënten konden op elk ogenblik beslissen met het onderzoek te stoppen zonder opgave van reden.

Behandelaren benaderden zelf alle cliënten uit hun caseload voor deelname aan het onderzoek naar de interbeoordelaarsbetrouwbaarheid en convergente validiteit van de FARE. Er golden twee inclusiecriteria: cliënten moesten 18 jaar of ouder zijn en de Nederlandse taal voldoende beheersen. Er waren verder geen redenen om cliënten te excluderen.

Cliënten die toestemming gaven voor deelname aan het deelonderzoek naar de predictieve validiteit van de FARE ontvingen geen vergoeding.

3.6 Analyses

Voor de statistische analyses voor alle studies is gebruik gemaakt van IBM SPSS Statistics versie 25. Per onderdeel van de studie staat bij het betreffende hoofdstuk beschreven welke statistische methoden zijn gebruikt voor het bestuderen van de onderzoeksvragen

4. Interbeoordelaarsbetrouwbaarheid

4.1 Methode

Onderzoeksgroepen

Van de 84 geïnterviewde cliënten, was 67,9% man ($n = 57$), 8,3% vrouw ($n = 7$), en voor 23,8% was het geslacht onbekend ($n = 20$). De leeftijd was bekend van 60 cliënten: deze waren gemiddeld 39,7 jaar ($SD = 12,6$; range 20-68). De helft (51,2%) van de cliënten was in behandeling bij de Waag, 20,2% bij Fivoor, 15,5% bij Transfore, 8,3% bij Kairos en 4,8% bij VNN. Bij 14,3% werd als primaire diagnose periodiek explosieve stoornis gesteld, bij 27,4% een stoornis in de impulscontrole of een gedragsstoornis en bij 9,5% (trekken van) een antisociale persoonlijkheidsstoornis. Verder kwamen ook de vermijdende persoonlijkheidsstoornis, borderline persoonlijkheidsstoornis, persoonlijkheidsstoornis NAO, pedofiele stoornis, stoornis in alcohol of drugsgebruik, obsessieve-compulsieve stoornis, posttraumatische stress stoornis en ADHD voor. Voor 27,4% was de primaire diagnose niet bekend. Vanwege strikte privacyregels binnen Kairos, Transfore, en Fivoor werden, ondanks schriftelijke toestemming van de cliënt, persoonsgegevens niet verstrekt.

Beoordelaars

Voor het berekenen van de interbeoordelaarsbetrouwbaarheid zijn twee typen beoordelaars betrokken in het onderzoek, namelijk behandelaren en studenten. De overeenstemming werd onderzocht tussen enerzijds behandelaren en studenten en anderzijds studenten onderling. Om behandelaren zoveel mogelijk te ontlasten werd er voor gekozen om studenten in te zetten als tweede onafhankelijke beoordelaars. In eerder onderzoek werden in risicotaxatie getrainde studenten ook gebruikt om risicotaxaties te verrichten en vergeleken met behandelaren kwamen zij tot vergelijkbare scores (Douglas, Ogloff, & Hart, 2003). Daarnaast hebben we ook studenten onderling vergeleken om potentiële afwijkingen in de overeenstemming tussen studenten en behandelaren te kunnen verklaren vanuit de samenstelling van koppels. Eerder onderzoek wees uit dat onderzoekers en hoofden behandeling niet verschillende qua overeenstemming, maar wel wat de betreft de interpretatie van scores (de Vogel & de Ruiter, 2003). De vergelijking tussen studenten en behandelaren heeft de beperking dat studenten minder ervaring hadden met de forensische populatie en het uitvoeren van risicotaxatie. De mate van overeenstemming tussen studenten zal een indicatie geven van de verklaring voor een eventuele slechte overeenstemming tussen behandelaren en studenten. Als de overeenstemming op een item vergelijkbaar is tussen de beoordelaarskoppels kan het resultaat worden toegeschreven aan de kwaliteiten van het instrument. Van de 21 betrokken behandelaren waren negen werkzaam bij de Waag, zeven bij Fivoor, drie bij VNN en twee bij Transfore. De behandelaren scoorden de FARE voor hun eigen cliënten zoals zij dat te doen gebruikelijk zijn, als onderdeel van de behandeling. Zij hebben daarom geen interviews van andere cliënten/behandelaren gescoord. De tweede beoordelaars waren negen studenten met minimaal een bachelor diploma, (forensische) psychologie (Universiteit van Amsterdam en Tilburg University) of een Masterdiploma Forensische Sociale Professional (Hogeschool Utrecht). Acht studenten waren voorafgaand aan het onderzoek gestart met een masteropleiding, een student startte na afloop van het onderzoek met een masteropleiding. Alle studenten hebben een training gevolgd in het gebruik van de FARE. De studenten liepen stage bij de Waag ($n = 4$), Fivoor ($n = 4$) en Transfore ($n = 1$) en werkten vanaf de stagelocatie aan het onderzoek onder begeleiding van een van de auteurs van dit rapport).

Interviewprotocol

Twee interviewprotocollen (kort en lang) zijn ontwikkeld door de auteurs van dit rapport. Het korte interviewprotocol is ontwikkeld om de 17 FARE-items op een gestructureerde wijze te kunnen uitvragen. Het korte interview bevat 52 vragen en het lange interview bevat daarnaast 48 instrument-specifieke vragen over het IFpBE, de LS/CMI en de RAF GGZ. Deze zijn gebruikt voor het onderzoek naar respectievelijk de interbeoordelaarsbetrouwbaarheid en de convergente validiteit (zie hoofdstuk 5). Beide interviewprotocollen zijn op te vragen bij de eerste auteur van dit rapport.

Procedure

Voor de afname van de interviews voor het onderzoek naar de interbeoordelaarsbetrouwbaarheid zijn alle betrokken behandelaren getraind in de FARE. De onderzoekers van de verschillende instellingen informeerden alle betrokken behandelaren over het doel van het onderzoek, de afname van het korte (en lange) interviewprotocol, de procedure voor uitleg aan en toestemming van de cliënt en de opname apparatuur.

Behandelaren namen de interviews op (met memorecorder) en scoorden daarna de FARE in het ROM-systeem van de eigen instelling. De interviews mochten op elk moment in de behandeling worden afgenomen, bij 93,5% van de cliënten werd het interview afgenomen bij de start van de behandeling, voor 4,3% betrof het een tussenmeting en voor 2,2% als eindmeting. De startmeting werd afgenomen na de intake en herhaalde metingen elke vier maanden (de Waag) of tenminste elk jaar (Fivoor, Kairos, Transfore en VNN). Elke student beluisterde de gehele interviewopname en vatte op basis van het interviewprotocol de opname samen en scoorde de instrumenten op basis van het interviewprotocol. Het samenvatten van de interviewopname diende als controlemiddel tijdens het scoren van de instrumenten en achteraf om per student na te gaan welke informatie meegenomen was in de scoring van de instrumenten. Bij de afname van de interviews is gestreefd naar het volledig uitvragen van alle informatie maar als de student de informatie uit de interviews onvoldoende vond om tot een score op een item te komen, werd het item niet ingevuld.

Per instelling werd door een onderzoeker van de instelling de persoonsgegevens (o.a. geslacht, leeftijd en opleidingsniveau) en de FARE (gescoord door de behandelaar) verzameld, geanonimiseerd en digitaal apart van het gecodeerde databestand bewaard.

Power

Om vast te stellen hoe groot de onderzoeksgroep moet zijn om de interbeoordelaarsbetrouwbaarheid te kunnen berekenen, is gebruik gemaakt van de Sample size calculator (web) van Arifin (2020; zie ook Bujang & Baharum, 2017). Hiervoor is de minimale acceptabele betrouwbaarheid (ρ_0) vastgesteld op $ICC = ,50$ en de verwachte betrouwbaarheid (ρ_1) $ICC = ,75$ (de gemiddelde ICC gevonden in Bogaerts, et al., 2018; Canton, et al., 2004; Hildebrand, et al., 2005; Rocque & Plubber-Beale, 2014; Schuringa, et al., 2014; Van Horn, et al., 2009). Bij twee beoordelaren per cliënt, een alpha van ,05 (one-tailed) en een Power (β) van ,80 is voor de Intra Class Correlatie Coëfficiënt (ICC) een aantal van 45 cliënten vereist.

Analyses

De overeenstemming tussen beoordelaars (student en student; student en behandelaar) werd berekend met Intraclass Correlation Coefficients (ICC's, Koo & Li, 2016). Omdat er sprake is van een random steekproef van beoordelaars uit een populatie is gekozen voor een two-way random model en absolute agreement. De 'single measures' zijn gerapporteerd omdat de FARE uiteindelijk door één beoordelaar gescoord gaat worden. De ICC is berekend op item niveau (de zes statische en 11 dynamische items) en op risico-inschatting niveau (actuarieel en klinisch gestructureerd). Voor de interpretatie van de resultaten zijn de richtlijnen van Koo en Li (2016) gebruikt, waarbij een $ICC < ,50$ als slecht wordt verondersteld; $,50 \leq ICC \leq ,74$ matig; $,75 \leq ICC \leq ,89$ goed en een $ICC \geq ,90$ uitstekend.

In de resultaten wordt gerapporteerd over verschillende aantallen beoordelaars-(paren). De 84 afgenomen interviews werden beoordeeld door zowel behandelaren als studenten. Maar door de beperkte stageduur van de studenten werden 42 interviews beoordeeld door een tweede student.

4.2 Resultaten

In Tabel 5 worden de ICC-waardes met het bijbehorende betrouwbaarheidsinterval (95%) en significantieniveaus gerapporteerd.

Student versus behandelaar

De interbeoordelaarsbetrouwbaarheid tussen student en behandelaar was voor acht van de 17 items goed tot uitstekend (47,1%), voor zes items matig (35,3%) en voor drie items slecht (17,7%). De overeenstemming was slecht op de dynamische risicofactoren beperkte vrijetijdsbesteding, instabiele woonsituatie en regelovertredend gedrag. Ook over de score op het klinisch ingeschat recidiverisico was de overeenstemming slecht. De overeenstemming op de statische risicofactoren was beter dan op de dynamische factoren, hetgeen tevens leidde tot goede

overeenstemming op de actuariële classificatie van het recidiverisico (optelsom van de statische factoren).

Student versus student

De interbeoordelaarsbetrouwbaarheid tussen studenten was voor vijf van de 17 items goed tot uitstekend (29,4%) en voor negen items matig (52,9%). Voor drie dynamische risicofactoren - beperkte vrijetijdsbesteding, disfunctionele oplossingsvaardigheden en regelovertrekend gedrag - was de overeenstemming slecht (17,7%). Ook tussen studenten was de overeenstemming op het actuariële ingeschat recidiverisico goed, en een matige overeenstemming op het klinisch-gestructureerd ingeschat recidiverisico.

Tabel 5. Overeenstemming tussen behandelaren en studenten op de FARE-items van versie 1

	Student - Behandelaar (N = 84)			Student - Student (N = 42)		
	ICC	.95 BI	n	ICC	.95 BI	n
Statische risicofactoren						
S1 Leeftijd eerste politiecontact	,883***	,821 - ,924	76	,942***	,895 - ,968	42
S2 Aantal eerdere en huidige veroordelingen	,874***	,809 - ,918	77	,947***	,904 - ,971	42
S3 Diversiteit (dreigend) delictgedrag	,828***	,743 - ,887	77	,700***	,507 - ,826	42
S4 Regelovertrekend gedrag verleden	,784***	,681 - ,857	77	,751***	,585 - ,857	42
S5 Instabiliteit opleiding/werk verleden	,813***	,721 - ,877	77	,737***	,564 - ,848	42
S6 Problematisch middelengebruik verleden	,906***	,855 - ,940	73	,822***	,694 - ,900	42
Dynamische risicofactoren						
D1 Disfunctioneren opleiding/werk	,553***	,351 - ,701	75	,709***	,518 - ,832	42
D2 Financieel wanbeleid	,738***	,615 - ,826	75	,711***	,521 - ,834	41
D3 Delinquent sociaal netwerk	,712***	,579 - ,809	74	,614***	,387 - ,771	42
D4 Beperkte vrijetijdsbesteding	,423***	,222 - ,591	75	,411***	,125 - ,634	41
D5 Problematische (ex-) partnerrelatie	,773***	,655 - ,853	74	,859***	,750 - ,922	41
D6 Instabiliteit woonsituatie	,494***	,298 - ,650	73	,625***	,399 - ,780	41
D7 Problematisch middelengebruik	,852***	,774 - ,905	73	,539***	,278 - ,725	41
D8 Gebrekkige impulsbeheersing	,700***	,535 - ,808	75	,612***	,384 - ,770	42
D9 Disfunctionele oplossingsvaardigheden	,598***	,430 - ,725	75	,461**	,182 - ,670	42
D10 Antisociale houding	,584***	,412 - ,716	75	,654***	,436 - ,798	42
D11 Regelovertrekend gedrag	,495***	,305 - ,648	75	,148**	-,148 - ,421	42
Inschatting recidiverisico						
Actuariële inschatting recidiverisico	,888***	,829 - ,928	75	,860***	,755 - ,922	42
Klinische inschatting recidiverisico	,479***	,259 - ,653	59	,626***	,392 - ,784	40

5. Convergente validiteit

5.1 Methode

Onderzoeksgroep

Bij 52 (62%) van de 84 geïnterviewde cliënten werd een lang interview afgenomen ten behoeve van de convergente validiteit. Van deze cliënten was 48,1% man en 11,5% vrouw en voor 40,4% was het geslacht onbekend. De gemiddelde leeftijd van 24 cliënten was 39,5 jaar ($SD = 13,7$; range = 20-68), 44,2% van de cliënten was in behandeling bij de Waag, 23,1% bij Transfore, 13,5% bij Kairos, 13,5% bij Fivoor en 5,8% bij VNN. De meest gestelde primaire diagnoses waren: stoornis in de impulscontrole/ gedragsstoornis (15,4%), periodiek explosieve stoornis (9,6%), antisociale persoonlijkheidsstoornis (7,7%), ADHD (5,8%), borderline persoonlijkheidsstoornis (3,8%). Verder werden de volgende diagnoses gesteld: posttraumatische stress stoornis, parafiele stoornis, posttraumatische stoornis en eens stoornis in het gebruik van opioïden.

Instrumenten

Om uitspraken te kunnen doen over de validiteit van een instrument wordt doorgaans de correlatie berekend met een gouden standaard, dat wil zeggen een instrument waarvan de validiteit al is vastgesteld. Voor de inschatting van het recidiverisico bestaat geen gouden standaard. Om toch iets te kunnen zeggen over de vraag of met de FARE vergelijkbare constructen wordt gemeten als met andere risicotaxatie-instrumenten, werd er voor gekozen om de samenhang van de FARE te berekenen met het IFpBE, de LS/CMI en RAF GGZ. Deze zijn de meest gebruikte risicotaxatie-instrumenten in Nederland voor het inschatten van het recidiverisico in het ambulante forensische veld (Van Horn, et al., 2016). In paragraaf 3.2 is een uitgebreide beschrijving gegeven van het IFpBE, de LS/CMI, de RAF GGZ.

Procedure

In dit deelonderzoek werd dezelfde procedure voor het instrueren van de behandelaren gevolgd als bij de interbeoordelaarsbetrouwbaarheid.

Alle behandelaren waren getraind en hadden ervaring met de RAF GGZ (de Waag), de LS/CMI (Kairos en Transfore), de HKT-r en/of het IFpBE (Fivoor). De behandelaren van VNN hadden ervaring met verschillende instrumenten zoals de HCR, HKT-R en de LS/CMI. Behandelaren kregen - indien nodig - voorafgaand aan het afnemen van de interviews een uitgebreide instructie over het IFpBE, de LS/CMI en/of RAF GGZ. De negen studenten die als tweede beoordelaar de interviews voor de interbeoordelaarsbetrouwbaarheid hebben gescoord, scoorden ook de 52 lange interviews ten behoeve van de convergente validiteit. De studenten werden bij de start van het onderzoek getraind in het scoren van het IFpBE, de LS/CMI, de RAF GGZ en de FARE.

Power analyse

Voor het berekenen van de convergente validiteit is uitgegaan van een gemiddelde effectgrootte van ,40 (Jung, Ledi, & Daniels, 2013; Schuringa, et al., 2016) en een alpha van ,05 (one-tailed). Voor de convergente validiteit is een $N = 37$ vereist voor een power (β) van ,80 (Faul, Erdfelder, Lang, & Buchner, 2007; zie ook Cohen, 1988).

Analyses

Voor de convergente validiteit is de overeenstemming tussen de FARE-items en de items van het IFpBE, de LS/CMI en RAF GGZ onderzocht. In Tabel 6 is een overzicht opgenomen van de items uit de verschillende instrumenten die inhoudelijk hetzelfde beogen te meten. In bijlage 1 is een uitgebreide toelichting opgenomen van de inhoudelijk overeenkomende items.

Voordat de convergente validiteit werd berekend, zijn de items van de drie vergelijkbare instrumenten bewerkt. Als uitgangspunt gold de scoringswijze van de FARE; items worden allemaal in dezelfde richting gescoord, oplopend in ernst. De items uit de andere instrumenten werden - indien nodig - gespiegeld zodat bij deze items de betekenis van de scores eveneens oploopt in ernst.

In sommige risicotaxatie-instrumenten werden meer items gebruikt om eenzelfde construct te meten dan in de FARE (zie Tabel 6). Bijvoorbeeld, de LS/CMI gebruikt twee items (22. Geen recente deelname aan clubs/organisaties en 23. Mate van actieve vrijetijdsbesteding) om de

vrijtijdsbesteding in kaart te brengen. In deze gevallen werden de scores opgeteld. Deze somscore werd vervolgens afgezet tegen het inhoudelijk vergelijkbare FARE-item.

Het volledige overzicht van deze bewerkingen kan worden opgevraagd bij de eerste auteur.

Naast een vergelijking van items zijn ook de totaalscores van de instrumenten met de totaalscore van de FARE met elkaar vergeleken. Hiertoe zijn voor de FARE de scores op de statische en dynamische items bij elkaar opgeteld. Voor de LSCMI is de totaalscore op Sectie 1 gebruikt; voor de RAF GGZ zijn alle statische en dynamische items opgeteld. Voor het IFpBE zijn alle dynamische items opgeteld en vergeleken met de totaalscore van de dynamische items van de FARE omdat het IFpBE geen statische items bevat.

De convergente validiteit tussen de FARE en het IFpBE, de LS/CMI en de RAF GGZ is berekend met de Spearman Rho's of Pearson correlatiecoëfficiënten (afhankelijk van het meetniveau en de scoreverdeling). Voor de interpretatie van de correlatiecoëfficiënten is richtlijn van Cohen (1988) gehanteerd: $r_{(s)} < ,30$ is een zwakke samenhang, $,30 \leq r_{(s)} \leq ,50$ is een matige samenhang; $r_{(s)} > ,50$ is een sterke samenhang.

Tabel 6. Overzicht FARE-items en overeenkomende items van het IFpBE, de LS/CMI, en de RAF GGZ

FARE	IFpBE	LS/CMI (Sectie 1)	RAF GGZ Domein/item
Statische risicofactoren			
S1. Leeftijd eerste politiecontact	-	5	D1/f
S2. Aantal eerdere en huidige veroordelingen	-	1-3	D1/a
S3. Diversiteit delictgedrag	-	-	D1/abc
S4. Regelovertredend gedrag verleden	-	8	D1/h, D11/a
S5. Instabiliteit opleiding/werk verleden	-	10-12, 14	D2/abf
S6. Problematisch middelengebruik verleden	-	28, 29	D8/ab
Dynamische risicofactoren			
D1. Disfunctioneren opleiding/werk	6	9, 15-17	D2/deg hij
D2. Financieel wanbeleid	9, 10	43a	D3/ab
D3. Delinquent sociaal netwerk	17	21, 24-26, 27	D6/b
D4. Beperkte vrijetijdsbesteding	5	22, 23	D7/ab
D5. Problematische (ex-)partnerrelatie	20	18	D5/a
D6. Instabiliteit woonsituatie		43b	D4/ab
D7. Problematisch middelengebruik	M1, M3, M4	30-35	D8/cdefg
D8. Gebrekkige impulsbeheersing	11	-	D9/e
D9. Disfunctionele oplossingsvaardigheden	4	-	D9/g
D10. Antisociale houding	3, 12, 13	36, 37, 40	D10/ab
D11. Regelovertredend gedrag	2, 16		
Inschatting recidiverisico			
Actuariële inschatting	-	Act. inschatting	-
Klinisch gestructureerde inschatting	-	-	Klinisch gestr. inschatting

5.2 Resultaten

De correlaties tussen de dynamische items van de FARE en inhoudelijk corresponderende items van het IFpBE, de LS/CMI en de RAF GGZ waren matig ($r_s = ,30$) tot sterk ($r_s = ,87$). De correlaties tussen de statische items waren over het algemeen sterk (zie Tabel 7). Omdat het IFpBE geen statische items bevat, ontbreken de correlaties. De actuariële risico-inschatting van de FARE werd vergeleken met de actuariële risico inschatting van de LS/CMI en de klinische risico inschatting van de FARE werd vergeleken met de hoogste risico inschatting van de RAF GGZ. De overeenstemming tussen de items in beide vergelijkingen was sterk ($r_s = ,76$ en $r_s = ,81$). Ook de overeenstemming tussen de totaalscores van de FARE en de totaalscores van de drie instrumenten was sterk ($r = ,73$ en $r = ,77$). De zes – overwegend dynamische - FARE-items die matig correleerde met één of meerdere instrumenten worden hieronder besproken.

Regelovertrekend gedrag - Statisch 4 (S4) en Dynamisch 11 (D11). Zowel het statische als het dynamische FARE-item Regelovertrekend gedrag kwam matig overeen met de items van de drie andere instrumenten. In het item Regelovertrekend gedrag moet het overtreeden van juridische maatregelen én de motivatie voor behandeling (de bereidheid tot gedragsverandering en houden aan behandelafspraken) gewogen worden in de score op het item. Hiermee lijkt in de FARE-items meer nadruk te liggen op motivatie voor behandeling dan in de items van het IFpBE, de LS/CMI en de RAF GGZ.

Disfunctioneren opleiding/werk (D1). Het FARE-item Disfunctioneren opleiding/werk komt matig overeen met LS/CMI-items, die naast omgang met medeleerlingen/collega's en docenten/leidinggevenden ook in kaart hoe mensen presteren. Dit laatste vormt geen onderdeel in het FARE item.

Delinquent sociaal netwerk (D3). Delinquent sociaal netwerk komt matig overeen met het IFpBE-item. Mogelijk komt dit doordat in het IFpBE naast het hebben van delinquente vrienden (zoals in het FARE-item) ook meegenomen wordt of deze vrienden een grote rol spelen in het leven van de cliënt.

Beperkte vrijetijdsbesteding (D4). Het FARE-item Vrijetijdsbesteding waarin gemeten wordt of iemand passieve en/of actieve vrijetijdsbestedingen heeft en deze alleen en/of met anderen onderneemt komt matig overeen met de items van het IFpBE en de LS/CMI. De vrijetijdsbesteding gemeten in het IFpBE meet alleen of iemand in staat is om een evenwichtige dagindeling te maken, daarmee wordt niet meegewogen wat iemand doet en of dat met anderen gedaan wordt. En het LS/CMI-item meet of iemand deelneemt aan groepsactiviteiten/clubs en in welk mate het een actieve vrijetijdsbesteding betreft, zoals sporten. De FARE geeft anders dan de LS/CMI ook een niet-problematische score als iemand een individuele of passieve vrijetijdsbestedingen heeft.

Problematische (ex-) partnerrelatie (D5). Tot slot, problematische (ex-)partnerrelatie komt matig overeen met het RAF GGZ-item, waarin enkel de relatie met de huidige partner beoordeeld wordt en niet de relatie met een ex-partner.

Tabel 7. Correlatiecoëfficiënten tussen de FARE en het IFpBE, de LS/CMI en de RAF GGZ

FARE-items		IFpBE		LS/CMI		RAF GGZ	
		r_s	n	r_s	n	r_s	n
Statische risicofactoren							
S1	Leeftijd eerste politiecontact	-	-	,652***	50	,831****	52
S2	Aantal eerdere/huidige veroordelingen	-	-	,924***	52	,648***	52
S3	Diversiteit delictgedrag	-	-	-	-	,790*****	50
S4	Regelovertrekend gedrag verleden	-	-	,452***	52	,472*****	52
S5	Instabiliteit opleiding/werk verleden	-	-	,656****	52	,609****	52
S6	Problematisch middelengebruik verleden	-	-	,845****	47	,773***	52
Dynamische risicofactoren							
D1	Disfunctioneren opleiding/werk	,672****	35	,326***	51	,508****	51
D2	Financieel wanbeleid	,678****	51	,802****	51	,612***	51
D3	Delinquent sociaal netwerk	,298*	49	,657****	49	,814***	49
D4	Beperkte vrijetijdsbesteding	,349**	49	,333**	50	,592*****	50
D5	Problematische (ex-)partnerrelatie	,747****	33	,612***	51	,424***	51
D6	Instabiliteit woonsituatie	-	-	-	-	,844***	50
D7	Problematisch middelengebruik	,648****	49	,828***	49	,874****	46
D8	Gebrekkige impulsbeheersing	,753***	51	-	-	,833***	50
D9	Disfunctionele oplossingsvaardigheden	,602***	51	-	-	,570****	51
D10	Antisociale houding	,562****	51	,799****	51	,751****	51
D11	Regelovertrekend gedrag	,464***	51	-	-	-	-
	Actuariële inschatting recidiverisico	-	-	,756***	52	-	-
	Klinische inschatting recidiverisico	-	-	-	-	,861***	50
	Totaalscore	,734***	52	,733****	52	,767****	52

- = geen overeenkomende items; ^a = Pearson correlatiecoëfficiënt; < ,30 = zwakke samenhang, ,30 -,50 = matige samenhang; > ,50 = sterke samenhang; *p < ,05; **p < ,01; *** p < ,001

6. Veranderbaarheid FARE

6.1 Methode

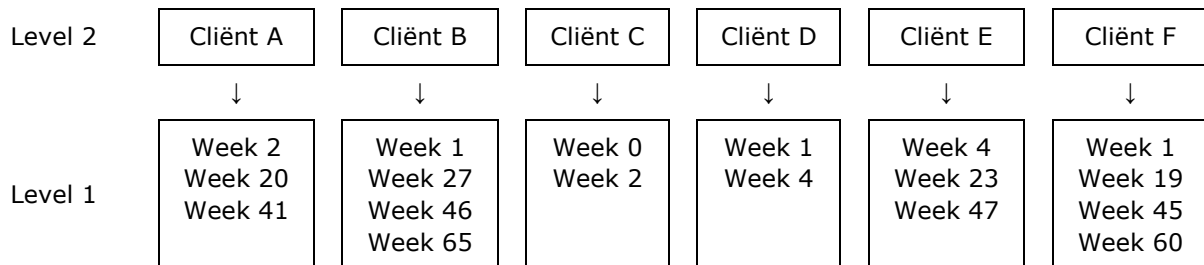
Onderzoeksgroep

De onderzoeksgroep voor het bestuderen van de sensitiviteit voor verandering van de FARE bestond uit 2180 cliënten die vanwege (dreigend) delictgedrag in de periode januari 2017 – januari 2018 gestart zijn met een behandeling bij de Waag, Transfore of Fivoor én toestemming hebben gegeven voor deelname aan wetenschappelijk onderzoek. De behandeling was bij de start van het onderzoek door $n = 451$ (20,7%) cliënten afgerond. De gemiddelde leeftijd van de groep was ten tijde van de intake 36,9 jaar (SD 12,4; range 17-83) en 88,6% was man. Van 446 cliënten waren twee metingen beschikbaar, van 870 cliënten drie metingen en van 864 cliënten vier metingen. De tijd tussen de metingen varieerde per cliënt.

Analyses

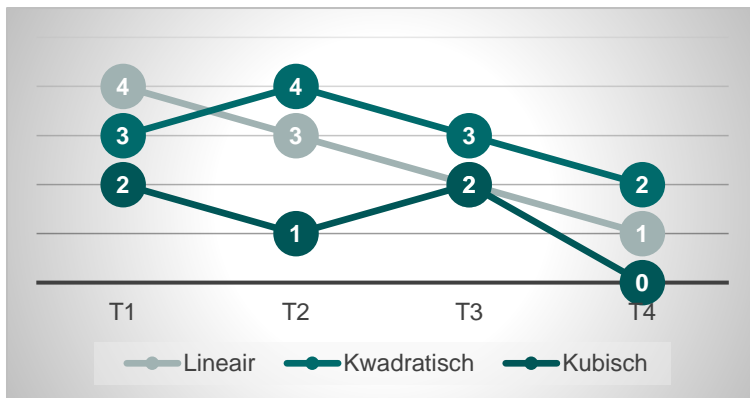
De veranderbaarheid van de FARE is getoetst met een Lineair Multilevel (groei-)Model (LMM) omdat per cliënt verschillende aantallen metingen kunnen worden meegenomen, de tijd tussen metingen mag variëren (zowel van een cliënt als tussen cliënten, zie Figuur 1). Met andere woorden: 1) de herhaalde FARE-metingen (level 1) zijn genest in de cliënten (level 2). Er is sprake van nesting omdat per cliënt, de verschillende metingen variëren in score maar niet volledig onafhankelijk van elkaar zijn. Tot slot is het bij multilevel groei modellen geen vereiste dat de data voldoen aan de assumptie van sphericiteit. Sphericiteit houdt in dat de variantie in de verschillen (tussen de meetmomenten) ruwweg gelijk is voor de groepen / cliënten (Heck, Thomas, & Tabata, 2014).

De LMM geeft de mogelijkheid om de verandering in de dynamische factoren over tijd te analyseren door de scores over de tijd te interpoleren (dat wil zeggen, de scores van metingen voor en na het ontbrekende meetmoment te middelen). Bijvoorbeeld: de verandering tussen T0 (de startscore) en T1 na vier maanden wordt vertaald naar een veranderscore per maand. En het verloop van de verandering kan worden geëxtrapoleerd. Voor de cliënten met drie meetmomenten wordt een voorspelling gedaan van de verandering van meetmoment drie naar meetmoment vier op basis van de cliënten met vier meetmomenten in de onderzoeksgroep.



Figuur 1: Multi-Level Model (getallen zijn voorbeelden)

Voor het toetsen van de veranderbaarheid van alle dynamische risicofactoren (afhankelijke variabelen) van de FARE is gebruik gemaakt van een Lineair Mixed Model (LMM) met een toegevoegde *higher order polynomial* met als inschatter maximum likelihood (ML). Lineair impliceert dat de verandering in een 'rechte lijn' verloopt, de klachten nemen in de loop van de tijd af, toe of blijven gelijk. Echter, de verwachting is dat de verandering op de risicofactoren fluctueert over de tijd. Door de *higher order polynomial* toe te voegen aan de LMM wordt niet alleen onderzocht of er een verandering is maar ook hoe deze verandering zich ontwikkelt. Er kan sprake zijn van een initiële stijging en daarna een daling of een stijging na een initiële daling, dit wordt een kwadratische verandering genoemd. En er is sprake van een S-vormige (kubische) verandering als een afname van risico (van startscore 2 op T1 naar score 0 op T2), wordt opgevolgd door een toename (score 2 op T3) om daarna weer af te nemen (score 0 op T4, Heck, et al., 2014). Zie in Figuur 2 de weergave van dit kubische verloop van de veranderingen.



Figuur 2. Voorbeeld groeimodel met Lineaire, kwadratische (stijging of daling) en kubische (S-vormige) verandering met score 0 – 4 op een FARE item.

De afhankelijke variabelen in de LMM zijn de 11 dynamische risicofactoren en het klinisch recidiverisico en de onafhankelijke variabele is 'tijd' in maanden (behandeling). Voor elke variabele werd in vier stappen een multilevel groeimodel opgesteld. In het nulmodel wordt de geschatte startscore bepaald op basis van de eerste meting van alle cliënten. De variabele 'tijd' wordt dus nog niet geïncorporeerd. In dit model wordt de Intraclass Correlatie Coëfficiënt (ICC), de overeenstemming tussen cliënten, bepaald. De ICC is gelijk aan de tussen-client variantie gedeeld door de totale variantie, wat resulteert in een waarde tussen '0' (alle observaties zijn onafhankelijk) en '1' (alle observaties van een groep zijn hetzelfde) (Sunhee & Lake, 2005). Hoge ICC-waarden ($\geq 0,25$) geven aan dat de data genest is (meerdere metingen per cliënt) en dat een multilevel aanpak nodig is om de data te analyseren. In Model 1 wordt de variabele 'tijd' toegevoegd om de mate van verandering in scores over een vastgestelde tijdseenheid te bepalen (Singer & Willet, 2003). Daarnaast wordt in dit model de geschatte startscore opnieuw bepaald met behulp van de eerste en tweede meting en wordt de mate van verandering over tijd bepaald (over alle cliënten met twee metingen). In Modellen 2 en 3 worden respectievelijk de Kwadratische (drie metingen) en kubische verandering (vier metingen) toegevoegd om het groeimodel te specificeren.

Om de modellen te vergelijken zijn de 'goodness of fit'-maten AIC (Akaike's Information Criterion) en BIC (Bayes Information Criterion) van elke stap vergeleken. Hoe kleiner de waarde van deze fitmaten, hoe beter het model bij de data past (Holden, Kelley, & Agarwal, 2008).

Reliable Change Index (RCI)

De veranderingen op de items van de FARE op cliëntniveau zijn berekend met de Reliable Change Index (RCI). RCIs worden berekend op basis van de startmetingen en eindmetingen/laatst afgenomen meting met de FARE met de formule $RCI = 1,96 * \sqrt{2 * SEM^2}$, waarbij $SEM = SD * \sqrt{(1 - r)}$. De 'r' verwijst naar een maat van betrouwbaarheid van het instrument; in deze studie is gebruik gemaakt van de interbeoordelaarsbetrouwbaarheid (ICC) tussen de behandelaar en de student (zie Tabel 5).

Als de RCI significant is, kan met 95% waarschijnlijkheid verondersteld worden dat de start en eindmeting van elkaar verschillen. In de huidige studie zijn de RCI-resultaten verdeeld in drie categorieën: betrouwbare verbetering, betrouwbare verslechtering en onveranderd. De eerste twee categorieën refereren aan scores die voldoen aan het criterium van $\pm 1,96$ significante verandering. De categorie 'onveranderd' bevat scores die wel een verbetering of verslechtering laten zien maar geen betrouwbare verandering.

De referentiewaarden van Cohen (1988) zijn gebruikt voor de interpretatie van de grootte van de verandering: klein $,20 \geq d \geq ,49$, matig $,50 \geq d \geq ,79$ groot $d \geq ,80$.

Missing values

De LMM verwijderd alle cases met een ontbrekende waarde uit de analyse. De percentages ontbrekende gegevens waren zeer laag, tussen 0% en 1,1% voor D2 'Financieel wanbeleid' van de eerste meting. Het effect van deze ontbrekende waarden op de resultaten is verwaarloosbaar waar later op wordt ingegaan. De ontbrekende waarden zijn daarom opgevuld met de *expectation-maximization algorithm* (EM). Een voorwaarde voor het gebruik van EM is de veronderstelling dat de ontbrekende waarden alleen toegewezen kunnen worden aan de geobserveerde waarden (Missing At Random, MAR of Missing Completely at Random MCAR). Om dit vast te stellen werd de

Little's MCAR test uitgevoerd. Ondanks het zeer lage percentage ontbrekende waarden, was Little's MCAR test significant. Little's MCAR test is gevoelig voor de grootte van de onderzoeksgroep (N): hoe hoger de N, des te groter de kans op significante (vals-positieve) resultaten. De ontbrekende gegevens zijn op *face value* onderzocht. Er werd geen samenhang gevonden tussen de ontbrekende informatie en een specifieke groep binnen de cliënten of de beoordelaars. De ontbrekende waarden zijn Missing at Random waardoor het mogelijk is om de ontbrekende waarden op te vullen met EM zonder dat het een vertekening geeft van de onderzoeksresultaten.

Kanskapitalisatie

Kanskapitalisatie is een statistisch verschijnsel dat zich voordoet wanneer een onderzoeker een reeks inferentiële toetsen uitvoert op basis van hetzelfde waarnemingsmateriaal. Het probleem bij kanskapitalisatie is dat het significantieniveau α (alfa) stijgt bij het uitvoeren van een reeks toetsen. Naarmate er meer variabelen worden gebruikt om gegevens te modelleren, stijgt de kans dat er een significant verband tussen twee van die variabelen wordt gevonden, terwijl dat verband op toeval berust.

De Bonferroni correctie beschermt tegen de kans op type-1 fouten (vertaald naar het huidige onderzoek: het ten onrechte aannemen dat een voorspeller een significante bijdrage levert aan het voorspellen van recidive). Om de kans op dergelijke type I fouten te verkleinen wordt de p -waarde aangepast aan het aantal toetsen (Armstrong, 2014). Voor de LMM is de aangepaste waarde $p = ,004$ ($p = ,05 / 12$ items). Het corrigeren voor type-I fouten werkt type-II fouten in de hand (Armstrong, 2014), namelijk de kans neemt toe dat ten onrechte een voorspellende variabele als niet-significant wordt bestempeld. Om te waken voor overcorrectie op type-I fouten wordt daarom de significante-waarde gedeeld door het aantal toetsen per subgroep (12 items x 2 groepen of 21 items x 8 subgroepen) in plaats van alle toetsen over alle subgroepen (respectievelijk 24 en 168 toetsen).

6.2 Resultaten

Voor de gehele onderzoeksgroep van 2180 cliënten werd per dynamische risicofactor en de klinische inschatting van het recidiverisico onderzocht of er sprake was van verandering over tijd, en zo ja, hoe deze verandering er uit ziet (rechtlijnig of meer gekromd). De verandering werd gemeten met minimaal twee tot maximaal vier FARE-metingen.

Er werden vier zogenaamde groei modellen getoetst met als basis het nul-model. Het eerste groei model is het nul-model met enkel de startscore (model 0). Het nul-model bestaat uit de ICC-waardes die werden berekend door de tussen-client variantie te delen door de totale variantie. Hiervoor werden de data van de eerste meting gebruikt, ook wel de startscore genoemd. De gemiddelde ICC-waarde was ,70 ($SD = ,10$; range ,49 - ,82). Vervolgens werd onderzocht of er significante veranderingen in scores waren gedurende de behandeling en hoe deze verandering er uit ziet: groei model 1 (is de verandering lineair?), groei model 2 (is de verandering kwadratisch?) en groei model 3 (is de verandering kubisch of te wel S-vormig?). Het tweede model (model 1) wordt ook wel lineair model genoemd en de modellen 2 en 3 worden ook wel de *higher order polynomial* modellen genoemd.

Anders gezegd wordt in vier stappen getoetst welk groei model het beste de verandering in een risicofactor beschrijft. Het model beschrijft de verandering van een risicofactor voor de ambulante populatie als elk onderdeel (startscore en/of lineaire verandering en/of kwadratische verandering en/of kubische verandering) een significante bijdrage levert in de beschrijving van het groei model.

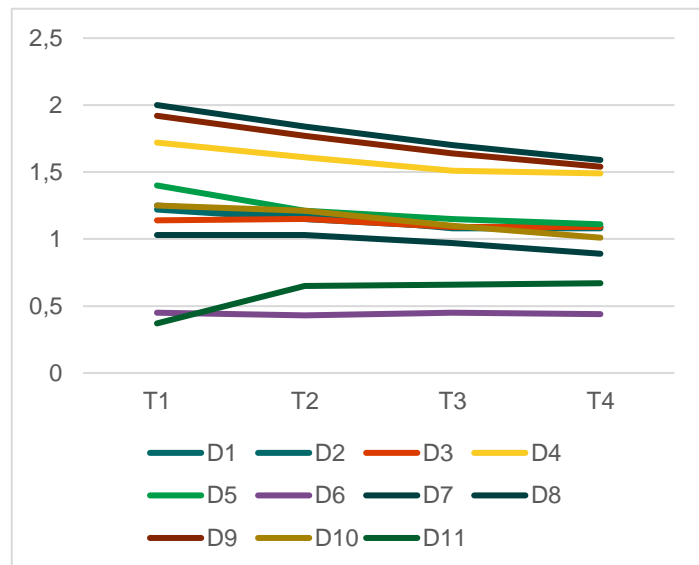
De LMM-analyse is uitgevoerd voor de totale groep (tabel 8) en voor de groep cliënten met een hoge score op de risicofactoren (tabel 9). In Tabellen 8 en 9 zijn voor het leesgemak het groei model dat de verandering van een risicofactor het beste beschrijft dikgedrukt. Daarnaast zijn vier grafieken (figuren 3 – 6) met gemiddelde scores per meetmoment opgenomen die het verloop van de verandering illustreren. Figuur 3 geeft de veranderingen per item weer voor de totale groep en figuren 4 – 6 weerspiegelen de veranderingen voor de hoogscoorders.

In Tabel 8 wordt een samenvatting gegeven van de resultaten van het lineaire multilevel model (model 1) en de toegevoegde *higher order polynomial* (model 2 en 3) voor de totale groep. Het model waar alle elementen een significante bijdrage leveren met daarbij ook de kleinste AIC en BIC model-fitmaat beschrijft de verandering op de risicofactor het beste. Per risicofactor is het model dat de verandering het beste beschrijft dikgedrukt.

Voor alle risicofactoren en de gestructureerd klinische inschatting van het recidiverisico geldt dat model 1, 2 of 3, de groeicurve het beste beschrijft. Met andere woorden, er is altijd sprake van verandering over tijd. Maar hoe de verandering er uit ziet varieert: lineair, kwadratisch of kubisch (S-vormig). Hierna volgt een toelichting op de resultaten.

Lineaire verandering (model 1).

Risicofactor D6 Instabiliteit woonsituatie heeft de laagste startscore en er is geen sprake van verandering over tijd. Verandering op de risicofactoren D2 Financieel wanbeleid en D7 Problematisch middelengebruik laten een lineaire verbetering zien. Dit betekent dat met elke maand extra in behandeling de scores op deze risicofactoren afnemen met respectievelijk ,013 en ,007.



Figuur 3. Totale groep: per FARE-item de gemiddelde score op vier meetmomenten

Kwadratische verandering (model 2).

Het kwadratische groeimodel beschrijft het beste de verandering op de risicofactoren: D1 Disfunctioneren opleiding/werk, D3 Delinquent sociaal netwerk, D4 beperkte vrijetijdsbesteding en D10 Antisociale houding. De initiële verbetering op deze risicofactoren vlt in de loop van de behandeling af.

Kubische verandering (model 3). De risicofactoren D5 Problematische (ex-)partnerrelatie, D8 Gebrekkige impulsbeheersing, 9 Disfunctionele oplossingsvaardigheden, D11 Regelovertredend gedrag en de klinische inschatting van het recidiverisico laten de meest complexe geschatte (kubische) verandering over tijd zien: de score wijzigt na de start doorgaans driemaal van richting. Voor deze factoren is in eerste instantie een verbetering zichtbaar, waarna een verslechtering volgt om uiteindelijk weer (licht) te verbeteren. De risicofactor D11 Regelovertredend gedrag is hier een uitzondering op: de startscore ,633 (een score tussen 0 'cliënt is zeer gemotiveerd voor behandeling' en score 1 'cliënt houdt zich om externe redenen aan de behandelafspraken') zal in de loop van de behandeling toenemen. Dit betekent dat de cliënt minder gemotiveerd raakt voor behandeling. In de loop van de behandeling (meetmoment 3) zal de motivatie van de cliënt weer iets verbeteren.

Opvallend is dat de startscores voor alle risicofactoren en de gestructureerde klinische inschatting in de groeimodellen een risicoscore hebben lager dan 2. Op een 5-puntsschaal waar de maximale score vier is, kan een score van 2 of lager opgevat worden als het minimaal of gemiddeld aanwezig zijn van een risicofactor. De resultaten zijn echter gebaseerd op de totale groep waarbij per risicofactor de hoogscoorders gemiddeld worden met de laagscoorders. Anders gezegd, aannemelijk is dat hoogscoorders meer verandering laten zien gedurende de behandeling, maar dat dit niet zichtbaar wordt omdat zij samen zijn geanalyseerd met de laagscoorders, die minder (kunnen) veranderen. Om deze veronderstelling zichtbaar te maken is een tweede reeks LMM analyses uitgevoerd bij de groep hoog scorende cliënten (startscore 2 of hoger).

Tabel 8. Start scores (Fixed Intercepts) op FARE Risicofactoren van de gehele onderzoeksgroep $N = 2.180$ (modellen 0 en 1), $n = 1.743$ (model 2) en $n = 864$ (model 3)

Risicofactor	Model	Fixed Intercept (Est.)	SE	Rate of change (p/m)	SE	Kwadratisch Rate of change	SE	Kubisch Rate of change	SE	ICC	AIC	BIC
D1 Disfunctioneren opleiding/werk	0	1,170***	,023							,74	18.599	18.620
	1	1,169***	,023	-,009***	,001						18.539	18.567
	2	1,158***	,023	-,011***	,001	,0001***	,000				18.509	18.544
	3	1,148***	,024	-,012***	,001	,0003**	,000	-,000 ^{ns}	,000		18.508	18.550
D2 Financieel wanbeleid	0	1,139***	,023							,78	17.752	17.773
	1	1,137***	,023	-,013***	,001						17.608	17.636
	2	1,133***	,023	-,014***	,001	,000 ^{ns}	,000				17.605	17.640
	3	1,132***	,024	-,014***	,001	,000 ^{ns}	,000	-,000 ^{ns}	,000		17.607	17.649
D3 Delinquent sociaal netwerk	0	1,123***	,019							,81	13.720	13.741
	1	1,123***	,019	-,002**	,001						13.712	13.740
	2	1,117***	,019	-,003***	,001	,000**	,000				13.703	13.738
	3	1,126***	,019	-,002**	,001	-,000 ^{ns}	,000	,000 ^{ns}	,000		13.699	13.740
D4 Beperkt vrijetijdsbesteding	0	1,603***	,024							,66	19.210	19.231
	1	1,600***	,024	-,014***	,001						19.062	19.090
	2	1,593***	,024	-,016***	,001	,000**	,000				19.050	19.092
	3	1,584***	,025	-,016***	,001	,000 ^{ns}	,000	-,000 ^{ns}	,000		18.309	18.364
D5 Problematische (ex-)partnerrelatie	0	1,210***	,024							,74	21.150	21.171
	1	1,204***	,025	-,022***	,001						20.909	20.937
	2	1,188***	,024	-,026***	,001	,0002***	,000				20.854	20.889
	3	1,172***	,025	-,027***	,002	,0004***	,000	-,000**	,000		20.848	20.890
D6 Instabiliteit woonsituatie	0	,438***	,015							,64	13.858	13.879
	1	,437***	,015	-,002 ^{ns}	,001						13.853	13.881
	2	,435***	,015	-,003 ^{ns}	,001	,000 ^{ns}	,000				13.852	13.887
	3	,438***	,015	-,002 ^{ns}	,001	-,000 ^{ns}	,000	,000 ^{ns}	,000		13.854	13.896

Noot. Est. = schatting; ** $p \leq ,004$, *** $p < ,001$, ns = niet significant; Intercept = de geschatte startscores; Rate of Change = de mate van verandering; (p/m) = per maand op een schaal van 0 - 4; ICC = Intraclasse Correlatie Coëfficiënt; AIC = Akaike's Information Criterion; BIC = Bayes Information Criterion

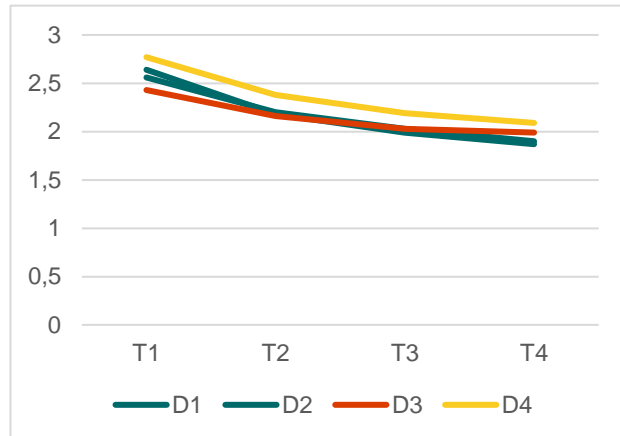
Tabel 8. Vervolg

Risicofactor	Model	Fixed Intercept (Est.)	SE	Rate of change (p/m)	SE	Kwadratisch Rate of change	SE	Kubisch Rate of change	SE	ICC	AIC	BIC
D7 Problematisch middelengebruik	0	1,008***	,026							,82	18.017	18.038
	1	1,007***	,026	-,005***	,001						17.999	18.027
	2	1,003***	,026	-,005***	,001	,000 ^{ns}	,000				17.997	18.032
	3	1,011***	,027	-,005***	,001	-,000 ^{ns}	,000	,000 ^{ns}	,000		17.995	18.043
D8 Gebrekkige Impulsbeheersing	0	1,807***	,019							,66	17.938	17.958
	1	1,803***	,020	-,023***	,001						17.531	17.559
	2	1,785***	,020	-,027***	,001	,0002***	,000				17.424	17.458
	3	1,764***	,020	-,029***	,001	,0005***	,000	-,000***	,000		17.405	17.446
D9 Disfunctionele oplossingsvaardigheden	0	1,750***	,016							,58	16.408	16.429
	1	1,747***	,016	-,018***	,001						16.077	16.109
	2	1,733***	,016	-,023***	,001	,0001***	,000				15.987	16.021
	3	1,717***	,016	-,024***	,001	,0004***	,000	-,000***	,000		15.972	16.019
D10 Antisociale houding	0	1,182***	,019							,77	15.151	15.172
	1	1,180***	,020	-,009***	,001						15.041	15.069
	2	1,172***	,019	-,011***	,001	,000***	,000				15.019	15.053
	3	1,171**	,020	-,011***	,001	,000 ^{ns}	,000	-,000 ^{ns}	,000		15.020	15.062
D11 Regelovertredend gedrag	0	,582***	,017							,49	18.973	18.993
	1	,585***	,017	,014***	,001						18.851	18.878
	2	,597***	,017	,018***	,001	-,0001***	,000				18.797	18.832
	3	,633***	,018	,022***	,001	-,0007***	,000	,000***	,000		18.738	18.779
Klinische risico-inschatting	0	1,893***	,020							,71	15.273	15.294
	1	1,885***	,020	-,025***	,001						14.737	14.765
	2	1,868***	,020	-,029***	,001	,0002***	,000				14.627	14.661
	3	1,854***	,020	-,030***	,001	,0004***	,000	-,000**	,000		14.619	14.660

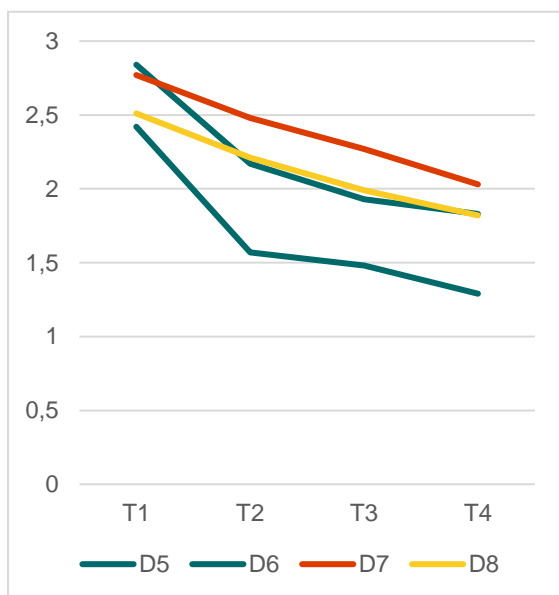
Noot. *Est.* = schatting; ** $p \leq ,004$, *** $p < ,001$, *ns* = niet significant; Intercept = de geschatte startcores; Rate of Change = de mate van verandering; (p/m) = per maand op een schaal van 0 - 4; ICC = Intraclasse Correlatie Coëfficiënt; AIC = Akaike's Information Criterion; BIC = Bayes Information Criterion

Tabel 9 beschrijft de groeicurves voor de groep met scores op de FARE-items van twee of hoger (bepaald per risicofactor) en de figuren 4 – 6 het verloop van de verandering. De gemiddelde ICC was ,44 (SD ,08; Range ,31 - ,53). De startscores voor deze subgroep liggen rond de 2 en de verandering in model 1 is per maand groter dan voor de gehele onderzoeksgroep (Zie Tabel 8). Risicofactor D2 Financieel wanbeleid heeft een startscore van 2,118 en de verandering is ten opzichte van het begin een verbetering (-,051) en vervolgens een lichte verslechtering (,001). Voor alle andere risicofactoren en de inschatting van het recidiverisico geldt dat een S-vormig (kubisch) groeimodel de verandering het beste beschrijft en na de lichte verslechtering de verandering weer richting verbetering gaat.

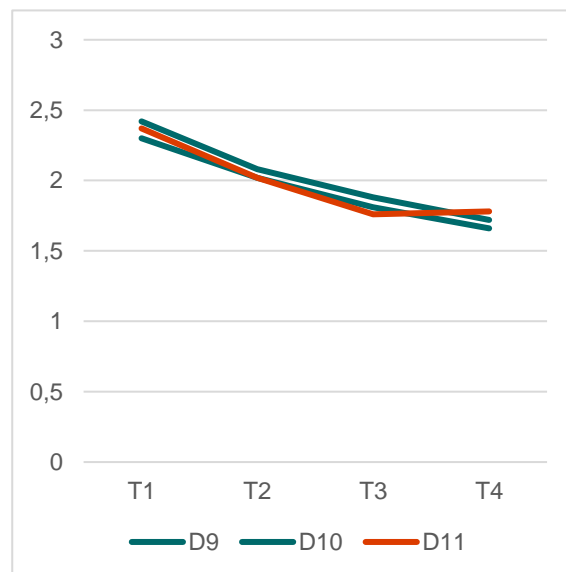
De resultaten van beide reeksen van LMM-analyses (tabel 8 en 9) samengevat geldt voor de gehele onderzoeksgroep en de subgroep met een problematische startscore dat er voornamelijk kwadratische en kubische veranderingen zijn op de risicofactoren. De veranderingen na de initiële (sterke) verbetering zijn echter klein ongeacht of het gaat om verbeteringen of een verslechtering.



Figuur 4. Hoogscorders: Gemiddelde scores op vier meetmomenten op domein 1 – 4



Figuur 5. Hoogscorders: Gemiddelde scores op vier meetmomenten op domein 5 – 8



Figuur 6. Hoogscorders: Gemiddelde scores op vier meetmomenten op domein 9 – 11

Tabel 9. Start scores (Fixed Intercepts) op FARE Risicofactoren van de groep met problematische startscore 2 of hoger

Risicofactor	Model	<i>n</i>	Fixed Intercept (Est.)	SE	Rate of change (p/m)	SE	Kwadratisch Rate of change	SE	Kubisch Rate of change	SE	ICC	AIC	BIC
D1 Disfunctioneren opleiding/werk	0	811	2,228***	,028							,45	7.471	7.489
	1	811	2,237***	,031	-,028***	,002						7.285	7.309
	2	626	2,207***	,029	-,043***	,002	,0002***	,000				7.115	7.144
	3	331	2,138***	,029	-,050***	,002	,001***	,000	-,000***	,000		7.035	7.071
D2 Financieel wanbeleid	0	880	2,198***	,025							,44	7.648	7.666
	1	880	2,170***	,025	-,045***	,002						7.157	7.181
	2	692	2,118***	,025	-,051***	,002	,001***	,000				7.088	7.119
	3	351	2,116***	,027	-,051***	,002	,001**	,000	,000 ^{ns}	,000		7.090	7.127
D3 Delinquent sociaal netwerk	0	620	2,183***	,026							,51	4.459	4.475
	1	620	2,185***	,028	-,017***	,002						4.367	4.390
	2	471	2,159***	,026	-,029***	,002	,0002***	,000				4.247	4.275
	3	226	2,129***	,027	-,032***	,002	,001***	,0001	-,000***	,000		4.232	4.266
D4 Beperkt vrijetijdsbesteding	0	1.112	2,395***	,025							,49	10.296	10.315
	1	1.112	2,384***	,025	-,039***	,002						9.870	9.895
	2	876	2,329***	,025	-,046***	,002	,001***	,000				9.790	9.822
	3	447	2,300***	,027	-,046***	,002	,002***	,0002	-,000**	,000		9.780	9.818
D5 Problematische (ex-)partnerrelatie	0	926	2,228***	,028							,32	10.161	10.179
	1	926	2,221***	,031	-,041***	,002						9.840	9.865
	2	780	2,179***	,029	-,064***	,002	,0003***	,000				9.596	9.627
	3	410	2,053***	,033	-,083***	,004	,002***	,0003	-,000***	,000		9.515	9.551
D6 Instabiliteit woonsituatie	0	215	1,742***	,053							,33	2.257	2.271
	1	215	1,707***	,055	-,058***	,004						2.105	2.124
	2	187	1,611***	,054	-,074***	,005	,001***	,000				2.053	2.076
	3	94	1,507***	,058	-,074***	,005	,004***	,001	-,000***	,000		2.034	2.062

Noot. Est. = schatting; ** $p \leq ,004$, *** $p < ,001$, ns = niet significant; Intercept = de geschatte startscores; Rate of Change = de mate van verandering; (p/m) = per maand op een schaal van 0 - 4; ICC = Intraclasse Correlatie Coëfficiënt; AIC = Akaike's Information Criterion; BIC = Bayes Information Criterion

Tabel 9. vervolg

Risicofactor	Model	n	Fixed Intercept (Est.)	SE	Rate of change (p/m)	SE	Kwadratisch Rate of change	SE	Kubisch Rate of change	SE	ICC	AIC	BIC
D7 Problematisch middelengebruik	0	714	2,465***	,032							,53	6.413	6.431
	1	714	2,458***	,034	-,027***	,002						6.248	6.271
	2	563	2,428***	,032	-,040***	,002	,000***	,000				6.151	6.180
	3	279	2,397***	,033	-,042***	,002	,001***	,000	-,000***	,000		6.141	6.176
D8 Gebrekkige Impulsbeheersing	0	1.567	2,179***	,017							,45	13.034	13.054
	1	1.567	2,173***	,019	-,030***	,001						12.528	12.555
	2	1.289	2,148***	,018	-,039***	,001	,000***	,000				12.282	12.315
	3	654	2,107***	,018	-,043***	,001	,001***	,000	-,000***	,000		12.201	12.240
D9 Disfunctionele oplossingsvaardigheden	0	1.459	2,075***	,016							,41	11.236	11.256
	1	1.459	2,068***	,017	-,029***	,001						10.691	10.718
	2	1.184	2,045***	,016	-,039***	,001	,000***	,000				10.439	10.472
	3	590	2,002***	,017	-,043***	,001	,001***	,000	-,000***	,000		10.348	10.387
D10 Antisociale houding	0	889	2,010***	,021							,47	6.322	6.340
	1	889	1,998***	,022	-,025***	,001						6.048	6.072
	2	693	1,971***	,021	-,036***	,002	,000***	,000				5.868	5.898
	3	330	1,934***	,022	-,039***	,002	,001***	,000	-,000***	,000		5.833	5.869
D11 Regelovertredend gedrag	0	235	2,021***	,053							,31	2.427	2.441
	1	235	2,016***	,052	-,039***	,005						2.364	2.382
	2	185	1,951***	,055	-,047***	,006	,001**	,000				2.356	2.380
	3	95	1,877***	,060	-,046***	,006	,003***	,001	-,000**	,000		2.349	2.377
Klinische risico-inschatting	0	1.432	2,236***	,018							,53	10.753	10.772
	1	1.432	2,214***	,020	-,032***	,001						10.128	10.155
	2	1.192	2,188***	,018	-,040***	,001	,000***	,000				9.889	9.921
	3	634	2,153***	,019	-,042***	,001	,001***	,000	-,000***	,000		9.844	9.883

Noot. *Est.* = schatting; ** $p \leq ,004$, *** $p < ,001$, *ns* = niet significant; Intercept = de geschatte startscores; Rate of Change = de mate van verandering; (p/m) = per maand op een schaal van 0 - 4; ICC = Intraclass Correlatie Coëfficiënt; AIC = Akaike's Information Criterion; BIC = Bayes Information Criterion

Reliable change index (RCI)

Nu we weten uit de LMM-analyses dat de scores op de risicofactoren in de loop van de behandeling veranderen, is het interessant om te achterhalen of deze veranderingen statisch betrouwbaar zijn. Om dat te kunnen bepalen, zijn de RCIs berekend. De resultaten van de RCIs worden beschreven voor twee groepen, de cliënten met een afgeronde behandeling ($n = 451$ Totaal afgerond) en de groep cliënten met een afgeronde behandeling en een hoge startscore (score ≥ 2 , n varieert tussen $n = 63$ voor D6 instabiliteit woonsituatie en $n = 317$ voor D8 Gebrekkige impulsbeheersing Hoogscorders afgerond).

Voor de totaal afgeronde groep blijkt dat de gemiddelde startscore laag is (startscore < 2) en het verschil met de gemiddelde eindscore klein. Over het algemeen geldt dat hoe kleiner het verschil is tussen de gemiddelde start- en eindscore, des te groter de verandering moet zijn om als statistisch betrouwbaar te worden beschouwd. Op individueel niveau is er sprake van een betrouwbare verbetering op de FARE-items als cliënten tussen de twee tot drie punten verbeteren. Dat is bijvoorbeeld het geval bij het item Problematisch middelengebruik (verbetering van 2 punten) en het item Beperkte vrijetijdsbesteding (verbetering van 3 punten) (zie Tabel 10).

Voor de groep cliënten met een hoge startscore (Hoogscorders afgerond, zie Tabel 11) is het verschil tussen de start- en eindmeting $0,39 \leq d \leq 0,98$. Cliënten met een hoge score op D7 problematisch middelengebruik maken een betrouwbare verbetering door als zij een punt lager scoren. Voor alle andere items geldt dat cliënten een betrouwbare verbetering doormaken als de score met twee punten is gedaald. Hoogscorders maken bij een relatief kleinere verandering al een statistische betrouwbare verandering/verbetering door.

Tabel 10. Resultaten RCI per FARE item ($N = 451$) bij cliënten met afgeronde behandeling

	Gemiddelde Startscore	<i>SD</i>	Gemiddelde Eindscore	<i>SD</i>	Start – Eind <i>d</i>	<i>n</i> hoge startscore	↓	-	↑	SEM	RCI
D1 Disfunctioneren opleiding/werk	1,29	1,26	1,18	1,21	0,09	162	3	441	7	0,85	2,35
D2 Financieel wanbeleid	1,43	1,31	1,31	1,25	0,09	219	6	423	22	0,67	1,85
D3 Delinquent sociaal netwerk	1,26	1,06	1,22	1,02	0,04	152	8	432	11	0,57	1,59
D4 Beperkte vrijetijdsbesteding	1,71	1,36	1,62	1,34	0,06	225	8	437	6	1,04	2,88
D5 Problematische (ex-)partnerrelatie	1,52	1,38	1,24	1,29	0,20	216	12	387	52	0,66	1,84
D6 Instabiliteit woonsituatie	0,59	0,92	0,53	0,86	0,07	63	9	426	16	0,66	1,83
D7 problematisch middelengebruik	1,16	1,38	1,04	1,31	0,09	160	6	420	25	0,53	1,48
D8 Gebrekkige impulsbeheersing	1,95	1,06	1,71	1,01	0,23	317	3	421	27	0,58	1,61
D9 Disfunctionele oplossingsvaardigheden	1,86	0,95	1,71	0,93	0,16	293	8	423	20	0,60	1,66
D10 Antisociale houding	1,41	1,15	1,29	1,07	0,10	215	2	446	3	0,74	2,06
D11 Regelovertredend gedrag	0,79	1,12	1,11	1,33	-0,29	105	40	400	11	0,79	2,20
FARE totaalscore	14,98	7,52	13,97	7,63	0,14	242	4	431	16	3,61	10,00

Noot. d = de maat voor grootte van de verandering; ↓ aantal cliënten met een betrouwbare verslechtering; - aantal cliënten zonder betrouwbare verandering; ↑ aantal cliënten met betrouwbare verbetering; SEM = standard error measurement; RCI = aantal punten op (schaal 0 – 4) dat een cliënt moet veranderen om te kunnen spreken van een statistisch betrouwbare verandering.

Tabel 11. Resultaten RCI per FARE item hoogscoorders met afgeronde behandeling

	n	Gemiddelde Startscore	SD	Gemiddelde Eindscore	SD	Start – Eind <i>d</i>	↓	-	↑	SEM	RCI
D1 Disfunctioneren opleiding/werk	162	2,70	0,76	2,33	1,07	0,50	1	143	18	0,51	1,40
D2 Financieel wanbeleid	219	2,62	0,76	2,32	0,95	0,40	1	196	22	0,39	1,07
D3 Delinquent sociaal netwerk	152	2,52	0,70	2,24	0,83	0,40	0	141	11	0,38	1,04
D4 Beperkte vrijtijdsbesteding	225	2,89	0,82	2,57	1,05	0,39	1	200	24	0,62	1,73
D5 Problematische (ex-)partnerrelatie	216	2,80	0,77	2,10	1,23	0,90	4	160	52	0,37	1,03
D6 Instabiliteit woonsituatie	63	2,51	0,78	1,75	1,28	0,98	2	45	16	0,56	1,54
D7 problematisch middelengebruik	160	2,85	0,76	2,38	1,18	0,61	3	112	45	0,30	0,82
D8 Gebrekkige impulsbeheersing	317	2,50	0,68	2,10	0,87	0,59	0	290	27	0,37	1,04
D9 Disfunctionele oplossingsvaardigheden	293	2,42	0,63	2,10	0,81	0,51	1	272	20	0,40	1,11
D10 Antisociale houding	215	2,44	0,68	2,08	0,87	0,53	0	196	19	0,44	1,22
D11 Regelovertredend gedrag	105	2,59	0,74	2,08	1,27	0,69	6	73	26	0,53	1,46

Noot. *d* = de maat voor grootte van de verandering; ↓ aantal cliënten met een betrouwbare verslechtering; - aantal cliënten zonder betrouwbare verandering; ↑ aantal cliënten met betrouwbare verbetering; SEM = standard error measurement; RCI = aantal punten op (schaal 0 – 4) dat een cliënt moet veranderen om te kunnen spreken van een statistisch betrouwbare verandering.

7. Predictieve validiteit

7.1 Methode

Onderzoeksgroep

De voorspellende waarde van de FARE is onderzocht op basis van de startmetingen (afgenomen binnen twee maanden na de start van de behandeling) bij een groep van 1268 cliënten die bij de Waag in behandeling zijn (geweest) vanwege (dreigend) delictgedrag. De groep bestond voor 87,2% uit mannen, 77% was geboren in Nederland en 22,7% buiten Nederland. De gemiddelde leeftijd was 36,7 jaar (*SD* 12,4; Range 18 – 83). De Waag wijst cliënten op basis van de aanmeldingsklacht toe aan een van drie delict-gerelateerde zorglijnen (Agressie en Vermogen (A&V), Huiselijk geweld betreffende elke vorm van geweld tegen personen binnen de relatie/familie (HG) en Seksueel grensoverschrijdend gedrag betreffende alle hands-on en/of hands-off delictgedragingen richting minder- en/of meerderjarige slachtoffers (SGG). Daarnaast heeft de Waag ook een jeugdigen (12-17 jaar) en jongvolwassenen (18-23 jaar) groep. In dit onderzoek was voor de subgroep Jongvolwassenen (cliënten tot en met 23 jaar) geen informatie over het delictgedrag beschikbaar daarom worden zij op basis van leeftijdsgroep beschreven. Daarnaast wordt hierna alleen gesproken over de groep 'Jongvolwassenen' omdat in dit onderzoek alleen cliënten van 18 jaar en ouder zijn onderzocht.

Van de totale groep van 1268 cliënten werd 41,5% behandeld voor het plegen van een agressief en/of vermogensdelict (zorglijn A&V), 37% voor huiselijk geweld (zorglijn HG), 10,4% voor seksueel grensoverschrijdend gedrag (zorglijn SGG) en 11,2% werd behandeld in de zorglijn Jongvolwassenen. De predictieve validiteit is onderzocht voor de gehele onderzoeksgroep en acht subgroepen (zie Tabel 12). De subgroepen zijn gevormd op basis van beschikbare informatie over de grootste cliëntgroepen die binnen de Waag behandeld worden en bij wie de FARE verplicht afgenomen moet worden.

De jongvolwassenen en de cliënten die vanwege seksueel grensoverschrijdend gedrag zijn aangemeld wijken het meeste af qua gemiddelde leeftijd ten opzichte van de overige subgroepen. De groep met seksueel grensoverschrijdend gedrag is gemiddeld het oudst. Vanzelfsprekend zijn de jongvolwassenen gemiddeld het jongst met 20,3 jaar omdat zij op leeftijd (tussen de 18 en 24 jaar) aan de zorglijn Jongvolwassenen worden toegewezen.

Tabel 12. Beschrijving subonderzoeksgroepen t.b.v. de predictieve validiteit

	<i>n</i>	Gemiddelde leeftijd	<i>SD</i>	Range
Geboren in NL				
Geboren in Nederland	288	36,3	12,8	18 – 83
Buiten Nederland geboren	976	38,4	10,8	18 – 68
Geslacht				
Mannen	1106	36,4	12,5	18 – 83
Vrouwen	162	37,2	11,8	18 – 70
Zorglijnen				
Agressie en vermogen	526	37,8	11,5	19 – 70
Huiselijk geweld	469	38,3	10,3	19 – 78
Seksueel grensoverschrijdend gedrag	131	44,5	14,5	18 – 83
Jongvolwassenen	142	20,3	14,5	18 – 24

Recidivecijfers

De Waag, Fivoor en Transfore hebben gegevens verzameld voor de predictieve validiteit van de FARE. De onderzoeksgroep die hier beschreven wordt, is een deel van de oorspronkelijk groep ($N = 2180$) waarover data verzameld is. In het onderzoeksvorstel stond dat recidivecijfers bij de Koninklijke Landelijke Politiedienst opgevraagd zouden worden omdat hiervoor een korte volgtijd zoals in dit onderzoek geen bezwaar zou zijn. Helaas waren zij niet in staat om de cijfers aan te leveren. Daarom zijn in het onderzoek de veroordelingcijfers van het OBDJ gebruikt. Het OBDJ heeft de beperking dat de volgtijd na de verzamelde data minimaal twee jaar moet zijn omdat de administratieve verwerkingstijd van het JD twee jaar is. Bij een kortere volgtijd zouden eventuele recidives nog niet geregistreerd kunnen zijn wat onvolledige onderzoeksresultaten oplevert. Dit had

als gevolg dat alleen de gegevens konden worden geanalyseerd die in de periode 1 januari 2017 tot en met 31 december 2017 zijn verzameld. Dit had tot gevolg dat enkel gegevens van Waagcliënten zijn meegenomen omdat de gegevens bij Fivoor en Transfore voor deze deelstudie na december 2017 zijn verzameld.

Algemene recidive werd gedefinieerd als een nieuwe registratie van een delict in het Onderzoek- en Beleidsdatabase Justitiële Documentatie (OBJD) dat (nog niet) heeft geleid tot een veroordeling. Deze database is een kopie van het Justitiële Documentatie Systeem (JDS) en bevat actuele en historische strafrechtelijke gegevens van alle natuurlijk- en rechtspersonen. Anders dan in het JDS zijn de gegevens in het OBJD gepseudonimiseerd, alle identificerende gegevens zijn verwijderd of versleuteld. Recidive werd geregistreerd als een dichotome variabele met '1' = recidive en '0' = geen recidive. De follow-up periode was maximaal 36 maanden vanaf de eerste meting met de FARE. De recidive-gegevens werden opgevraagd op 9 januari 2020. In het OBJD was de meest recente update op 4 juli 2019, de recidivegegevens van de onderzoeksgroep zijn daarom tot die datum meegenomen.

Analyses

De twee meest gebruikte analyses voor het berekenen van de predictieve validiteit zijn de Receiver Operator Curve (ROC uitgedrukt in de Area Under the Curve (AUC)) en de Cox regressie survivalanalyse. De ROC levert een (area under the) curve op, gebaseerd op de verhouding tussen het percentage correct voorspelde recidivisten (sensitiviteit) en het percentage correct voorspelde niet-recidivisten (specificiteit). Dit levert een AUC-waarde op tussen 0 en 1 (perfecte voorspelling) van (niet-) recidivisten waarbij $AUC = ,50$ gelijk is aan een voorspellende waarde op kans-niveau. De referentiewaarde voor het effect van de AUC zijn als volgt: een klein effect als $,56 \leq AUC < ,64$, $,64 \leq AUC < ,71$ matig en $AUC \geq ,71$ is een groot effect (Rice & Harris, 2005). Naast de ROC werd ook een reeks Cox regressie survivalanalyses uitgevoerd.

In de Cox regressieanalyse wordt met behulp een 'status'- variabele en een 'tijd'- variabele een regressiemodel geschat voor recidivesnelheid. Daarbij kan worden onderzocht wat het individuele effect is van variabelen (risicofactoren) in het model op de recidivesnelheid. De 'tijd'-variabele is het aantal dagen tussen de eerste meting en de recidive en in het geval dat iemand niet is gerecidiveerd, het aantal dagen tussen de eerste meting en het einde van de volgtijd (4 juli 2019). Het effect van risicofactoren op de recidivesnelheid wordt uitgedrukt in *Hazard ratios* (de exponent van de regressie coëfficiënt b) en daarbij wordt een 95% betrouwbaarheidsinterval gegeven. Een *hazard ratio* met een waarde boven de 1 geeft de factor aan waarmee de recidivesnelheid toeneemt met elk punt toename in de score op de risicofactor. Een hazard ratio met een waarde onder 1 geeft een afname van de recidivesnelheid aan voor elk punt toename in de score op een risicofactor. De Chi^2 -test geeft met een p -waarde de algemene statistische significantie van het regressiemodel (Nikulin & Wu, 2016).

Kanskapitalisatie

Zoals besproken in hoofdstuk 6 is ook in deze deelstudie een Bonferroni correctie toegepast. De aangepaste p -waarde is $p = ,0024$ ($p = ,05 / 21$ items, FARE-items en somscores)

7.2 Resultaten

Voor de gehele onderzoeksgroep was de gemiddelde periode tussen de eerste FARE-meting en 4 juli 2019 (einde volgtijd) 683 dagen (SD 78,7; Range 552 - 849) en het recidivepercentage was 17,1%.

In tabel 13 wordt de gemiddelde volgtijd en recidivepercentages voor de subgroepen beschreven.

Tabel 13. Gemiddelde volgtijd en recidivepercentage voor acht subgroepen

	Volgtijd in dagen	SD	Range	Recidive
Geboren in Nederland	681	78,9	552 – 849	16,6%
Buiten Nederland geboren	686	78,2	560 – 843	19,1%
Mannen	684	78,5	552 – 849	17,4%
Vrouwen	670	79,3	554 – 843	16,0%
Agressie en vermogen	683	77,7	552 – 849	19,4%
Huiselijk geweld	687	78,5	552 – 849	17,4%
Seksueel grensoverschrijdend gedrag	676	80,0	561 – 843	4,6%
Jongvolwassenen (tot en met 23 jaar)	684	79,6	559 – 849	28,2%

*Tijd tussen eerste meting en einde volgtijd in dagen

Receiver Operator Curve

De resultaten van de ROC staan weergegeven in Tabellen 14 (gehele onderzoeksgroep, en subgroepen op basis van geboorteland en geslacht) en 15 (subgroepen op basis van zorglijn en jongvolwassenen).

De voorspellende waarde van de FARE-items voor de gehele onderzoeksgroep is zwak tot matig ($,58 \geq AUC \leq ,68$), met uitzondering van de items S2 Aantal eerdere en huidige veroordelingen deze voorspelde algemene recidive sterk ($AUC = ,71$) en D5 Problematische partnerrelatie heeft geen voorspellende waarde ($AUC = ,50$). De somscore van alle FARE-items geeft een sterke voorspelling van algemene recidive ($AUC = ,72$). Bij de grootste subgroepen, mannen (87%) en personen geboren in Nederland (77%) zijn de resultaten vergelijkbaar met de gehele onderzoeksgroep.

Van alle de risicofactoren is S2 Aantal eerdere en huidige veroordelingen de enige risicofactor die voor alle subgroepen (met uitzondering van de SSG-groep) een matige tot sterke voorspeller met een AUC 's tussen de $,67$ en $,76$.

Voor de SGG-subgroep geldt dat geen enkele factor een significante voorspellende waarde heeft voor algemene recidive. En voor de subgroep vrouwen geldt dat alleen S2 Aantal eerdere en huidige veroordelingen ($AUC = ,73$) en de klinische inschatting van het recidiverisico ($AUC = ,74$) significante voorspellers zijn.

De risicofactor Problematische (ex-)partnerrelatie blijkt voor geen van de subgroepen een significante voorspeller te zijn.

De subgroepen: cliënten geboren buiten Nederland, Agressie en Vermogen en Huiselijk geweld tonen de items (met uitzondering van S2) slechte tot matige AUC 's (respectievelijk $,61 \geq AUC \leq ,65$; $,58 \geq AUC \leq ,69$; $,61 \geq AUC \leq ,70$).

Tot slot blijken de items criminele voorgeschiedenis (S1 – S4), Disfunctioneren opleiding/werk (S5 en D1), D3 Sociaal netwerk, D6 Instabiliteit woonsituatie en D11 Regelovertredend gedrag voor de groep jongvolwassenen matige tot sterke voorspellers ($,67 \geq AUC \leq ,80$) te zijn.

Tabel 14. Voorspellende waarde (AUC) van de FARE-items en somscores voor de gehele onderzoeksgroep en subgroepen (geboorteland, geslacht)

	Totaal N = 1268		GbNL n = 288		GiNL n = 976		Man n = 1106		Vrouw n = 162	
	AUC	BI	AUC	BI	AUC	BI	AUC	BI	AUC	BI
S1 Leeftijd eerste politiecontact	,64***	,60 - ,69	,63**	,55 - ,71	,66***	,61 - ,70	,66***	,62 - ,70	,53 ^{ns}	,41 - ,65
S2 Aantal eerdere en huidige veroordelingen	,71***	,67 - ,74	,73***	,66 - ,80	,70***	,65 - ,74	,70***	,66 - ,74	,73***	,64 - ,82
S3 Diversiteit (dreigend) delictgedrag	,66***	,62 - ,70	,65**	,57 - ,73	,66***	,62 - ,70	,67***	,63 - ,71	,59 ^{ns}	,49 - ,70
S4 Regelovertredend gedrag in het verleden	,63***	,59 - ,67	,64**	,55 - ,72	,62***	,57 - ,67	,64***	,59 - ,68	,54 ^{ns}	,42 - ,66
S5 Instabiliteit opleiding/werk in het verleden	,65***	,61 - ,69	,63**	,55 - ,71	,65***	,60 - ,70	,66***	,62 - ,70	,54 ^{ns}	,41 - ,66
S6 Problematisch middelengebruik in het verleden	,59***	,55 - ,64	,59 ^{ns}	,50 - ,68	,59***	,55 - ,64	,59***	,55 - ,64	,57 ^{ns}	,43 - ,70
D1 Disfunctioneren opleiding/werk	,63***	,59 - ,67	,59 ^{ns}	,50 - ,67	,64***	,59 - ,69	,64***	,59 - ,68	,55 ^{ns}	,43 - ,67
D2 Financieel wanbeleid	,62***	,58 - ,67	,62 ^{ns}	,53 - ,71	,62***	,57 - ,67	,63***	,58 - ,67	,56 ^{ns}	,43 - ,69
D3 Delinquent sociaal netwerk	,68***	,63 - ,72	,64**	,55 - ,72	,69***	,64 - ,73	,68***	,64 - ,73	,61 ^{ns}	,49 - ,73
D4 Beperkte vrijetijdsbesteding	,60***	,56 - ,64	,61 ^{ns}	,52 - ,69	,60***	,55 - ,64	,59***	,55 - ,64	,61 ^{ns}	,50 - ,72
D5 Problematische (ex-) partnerrelatie	,50 ^{ns}	,46 - ,55	,48 ^{ns}	,40 - ,56	,51 ^{ns}	,46 - ,56	,51 ^{ns}	,47 - ,56	,43 ^{ns}	,31 - ,56
D6 Instabiliteit woonsituatie	,64***	,59 - ,68	,55 ^{ns}	,47 - ,64	,67***	,62 - ,71	,65***	,61 - ,69	,53 ^{ns}	,40 - ,66
D7 Problematisch middelengebruik	,60***	,56 - ,64	,57 ^{ns}	,48 - ,66	,61***	,56 - ,65	,60***	,56 - ,65	,54 ^{ns}	,41 - ,67
D8 Gebrekkige impulsbeheersing	,60***	,55 - ,64	,58 ^{ns}	,49 - ,67	,60***	,55 - ,65	,59***	,55 - ,64	,61 ^{ns}	,49 - ,74
D9 Disfunctionele oplossingsvaardigheden	,58**	,53 - ,62	,57 ^{ns}	,48 - ,65	,58**	,53 - ,63	,57**	,53 - ,61	,61 ^{ns}	,50 - ,72
D10 Antisociale houding	,62***	,58 - ,66	,61**	,53 - ,69	,63***	,58 - ,68	,63***	,58 - ,67	,59 ^{ns}	,46 - ,72
D11 Regelovertredend gedrag	,61***	,57 - ,65	,51 ^{ns}	,43 - ,60	,64***	,59 - ,69	,62***	,57 - ,66	,55 ^{ns}	,42 - ,68
Klinische inschatting van het recidiverisico	,68***	,63 - ,73	,67**	,57 - ,78	,68***	,63 - ,74	,67***	,62 - ,73	,74**	,62 - ,85
FARE somscore	,72***	,68 - ,76	,69***	,61 - ,77	,73***	,69 - ,77	,73***	,69 - ,77	,65 ^{ns}	,54 - ,77
SOM FARE 6 statische factoren	,71***	,68 - ,75	,72***	,65 - ,79	,71***	,67 - ,75	,72***	,68 - ,76	,67 ^{ns}	,59 - ,76
SOM FARE 11 dynamische factoren	,69***	,65 - ,72	,64**	,56 - ,72	,70***	,66 - ,75	,70***	,66 - ,73	,62 ^{ns}	,49 - ,75

Ns = niet significant, ** $p \leq ,0024$, *** $p < ,001$. GbNL = cliënten geboren buiten Nederland en GiNL = cliënten geboren in Nederland. BI = 95% betrouwbaarheidsinterval.

Tabel 15. Voorspellende waarde (AUC) van de FARE-items en somscores voor de subgroepen (zorglijn, jongvolwassenen)

	Agressie & vermogen n = 526		Huiselijk geweld n = 469		SGG n = 131		JOVO n = 142	
	AUC	BI	AUC	BI	AUC	BI	AUC	BI
S1 Leeftijd eerste politiecontact	,57 ^{ns}	,50 - ,63	,64***	,57 - ,71	,67 ^{ns}	,41 - ,93	,73***	,64 - ,82
S2 Aantal eerdere en huidige veroordelingen	,67***	,62 - ,73	,70***	,63 - ,76	,55 ^{ns}	,30 - ,81	,76***	,68 - ,84
S3 Diversiteit (dreigend) delictgedrag	,60**	,54 - ,66	,67***	,60 - ,73	,72 ^{ns}	,52 - ,92	,68**	,59 - ,78
S4 Regelovertredend gedrag in het verleden	,60**	,54 - ,66	,62**	,54 - ,70	,60 ^{ns}	,35 - ,86	,68**	,58 - ,78
S5 Instabiliteit opleiding/werk in het verleden	,58 ^{ns}	,52 - ,64	,62**	,56 - ,69	,72 ^{ns}	,45 - ,99	,74***	,66 - ,83
S6 Problematisch middelengebruik in het verleden	,53 ^{ns}	,47 - ,59	,69***	,63 - ,76	,71 ^{ns}	,45 - ,96	,51 ^{ns}	,40 - ,61
D1 Disfunctioneren opleiding/werk	,57 ^{ns}	,51 - ,63	,63**	,55 - ,70	,59 ^{ns}	,37 - ,82	,68**	,59 - ,77
D2 Financieel wanbeleid	,63***	,57 - ,69	,56 ^{ns}	,48 - ,63	,60 ^{ns}	,37 - ,82	,66 ^{ns}	,56 - ,76
D3 Delinquent sociaal netwerk	,64***	,58 - ,70	,59 ^{ns}	,52 - ,67	,62 ^{ns}	,35 - ,89	,80***	,72 - ,87
D4 Beperkte vrijetijdsbesteding	,57 ^{ns}	,52 - ,63	,60 ^{ns}	,53 - ,67	,54 ^{ns}	,29 - ,78	,63 ^{ns}	,52 - ,74
D5 Problematische (ex-) partnerrelatie	,50 ^{ns}	,43 - ,56	,51 ^{ns}	,44 - ,59	,70 ^{ns}	,50 - ,90	,57 ^{ns}	,46 - ,68
D6 Instabiliteit woonsituatie	,61***	,55 - ,67	,61**	,54 - ,69	,75 ^{ns}	,53 - ,97	,67**	,58 - ,77
D7 Problematisch middelengebruik	,53 ^{ns}	,47 - ,60	,70***	,64 - ,77	,58 ^{ns}	,35 - ,80	,53 ^{ns}	,42 - ,64
D8 Gebrekkige impulsbeheersing	,57 ^{ns}	,50 - ,63	,61**	,54 - ,69	,55 ^{ns}	,31 - ,79	,56 ^{ns}	,46 - ,67
D9 Disfunctionele oplossingsvaardigheden	,57 ^{ns}	,51 - ,63	,59 ^{ns}	,51 - ,66	,49 ^{ns}	,23 - ,74	,51 ^{ns}	,41 - ,62
D10 Antisociale houding	,58 ^{ns}	,52 - ,64	,62**	,55 - ,69	,60 ^{ns}	,35 - ,85	,64 ^{ns}	,54 - ,74
D11 Regelovertredend gedrag	,61***	,55 - ,67	,53 ^{ns}	,45 - ,60	,58 ^{ns}	,31 - ,85	,71***	,62 - ,81
Klinische inschatting van het recidiverisico	,69***	,62 - ,76	,60 ^{ns}	,50 - ,69	,56 ^{ns}	,29 - ,82	,80***	,69 - ,91
FARE somscore	,66***	,60 - ,72	,71***	,65 - ,78	,70 ^{ns}	,40 - ,99	,78***	,70 - ,86
SOM FARE 6 statische factoren	,63***	,58 - ,69	,73***	,68 - ,79	,70 ^{ns}	,42 - ,99	,75***	,67 - ,84
SOM FARE 11 dynamische factoren	,65***	,59 - ,71	,66***	,60 - ,73	,64 ^{ns}	,38 - ,90	,73***	,65 - ,82

Ns = niet significant, ** $p \leq ,0024$, *** $p < ,001$; SGG = cliënten die in behandeling zijn vanwege Seksueel Grensoverschrijdend Gedrag en JOVO = jongvolwassen cliënten (18 – 24 jaar). BI = 95% betrouwbaarheidsinterval

Cox regressieanalyses

De uitkomsten van de Cox regressieanalyses geven zicht op welke FARE-items recidive voorspellen. De analyses laten ook zien of cliënten met een hoge score op een risicofactor sneller recidiveren dan cliënten met een lagere score.

De resultaten wijzen - na Bonferoni correctie ($p \leq ,0024$) - op een significant regressiemodel voor de totale groep (tabel 16) en voor bijna alle subgroepen, met uitzondering van de vrouwelijke cliënten (tabel 20) en cliënten in behandeling vanwege seksueel grensoverschrijdend gedrag (tabel 24). De significante resultaten geven aan dat het voorspellingsmodel met 17 FARE-items als geheel betrouwbaar onderscheid kunnen maken tussen cliënten die al dan niet recidiveren. In de totale groep en elke subgroep droegen bepaalde risicofactoren meer bij aan deze voorspelling dan andere. Welke risicofactoren dat zijn wordt in de tabellen 16 - 24 weergegeven. Samenvatting werden de volgende significante voorspellers gevonden.

Eerder en huidige veroordelingen (S2). Voor de gehele onderzoeksgroep (tabel 16), cliënten geboren in Nederland (tabel 18) en cliënten in behandeling vanwege Agressie en vermogensdelicten (tabel 21) voorspelt de statische risicofactor S2 Aantal eerdere en huidige veroordelingen algemene recidive.

Verder levert na de Bonferroni-correctie geen enkele statische risicofactor een individuele bijdrage aan de voorspelling van algemene recidive.

Dynamische risicofactoren. Alleen voor cliënten geboren in Nederland (tabel 18) zijn naast S2 Eerdere en huidige veroordelingen ook dynamische risicofactoren voorspellend, namelijk een instabiele woonsituatie en het vertonen van regelovertrekend gedrag in de afgelopen zes maanden, voorspelt dat zij sneller met een delict zullen recidiveren.

Voor jongvolwassenen (tabel 23) voorspelt alleen het hebben van delinquente vrienden ($HR = 2,301$) een verhoogde kans op recidive.

Tot slot voor cliënten geboren buiten Nederland (tabel 17), voor mannen (tabel 19) en cliënten in behandeling vanwege huiselijk geweld (tabel 22) geldt dat geen enkele risicofactor na de Bonferroni-correctie een individuele bijdrage levert aan de voorspelling van algemene recidive.

Tabel 16. Resultaten Cox regressieanalyse voor de gehele onderzoeksgroep N = 1268

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							164,896	17	,000
S1 Leeftijd eerste politiecontact	,082	,056	2,185	,139	1,086	,974 – 1,211			
S2 Aantal eerdere en huidige veroordelingen	,282	,066	18,184	,000	1,326	1,165 – 1,510			
S3 Diversiteit (dreigend) delictgedrag	,030	,084	,126	,723	1,030	,874 – 1,215			
S4 Regelovertredend gedrag in het verleden	-,011	,073	,023	,879	,989	,856 – 1,142			
S5 Instabiliteit opleiding/werk in het verleden	-,013	,073	,033	,855	,987	,855 – 1,139			
S6 Problematisch middelengebruik in het verleden	-,115	,079	2,138	,144	,891	,764 – 1,040			
D1 Disfunctioneren opleiding/werk	,071	,070	1,029	,310	1,073	,936 – 1,231			
D2 Financieel wanbeleid	,035	,060	,335	,563	1,035	,920 – 1,165			
D3 Delinquent sociaal netwerk	,203	,081	6,383	,012	1,226	1,047 – 1,435			
D4 Beperkte vrijetijdsbesteding	,056	,059	,874	,350	1,057	,941 – 1,188			
D5 Problematische (ex-) partnerrelatie	-,003	,050	,003	,957	,997	,903 – 1,101			
D6 Instabiliteit woonsituatie	,181	,065	7,769	,005	1,198	1,055 – 1,360			
D7 Problematisch middelengebruik	,111	,079	1,964	,161	1,117	,957 – 1,304			
D8 Gebrekkige impulsbeheersing	,031	,084	,134	,714	1,031	,875 – 1,214			
D9 Disfunctionele oplossingsvaardigheden	-,041	,101	,161	,688	,960	,788 – 1,171			
D10 Antisociale houding	-,057	,086	,437	,508	,945	,799 – 1,118			
D11 Regelovertredend gedrag	,136	,069	3,901	,048	1,146	1,001 – 1,312			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 17. Resultaten Cox regressieanalyse voor de groep geboren buiten Nederland n = 288

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							42,325	17	,000
S1 Leeftijd eerste politiecontact	-,004	,110	,001	,971	,996	,802 – 1,237			
S2 Aantal eerdere en huidige veroordelingen	,364	,137	7,028	,008	1,438	1,099 – 1,882			
S3 Diversiteit (dreigend) delictgedrag	,108	,178	,368	,544	1,114	,786 – 1,580			
S4 Regelovertredend gedrag in het verleden	,080	,162	,242	,622	1,083	,788 – 1,372			
S5 Instabiliteit opleiding/werk in het verleden	,008	,157	,003	,959	1,008	,741 – 1,372			
S6 Problematisch middelengebruik in het verleden	-,077	,169	,208	,648	,926	,665 – 1,289			
D1 Disfunctioneren opleiding/werk	-,089	,153	,336	,562	,915	,678 – 1,235			
D2 Financieel wanbeleid	,111	,117	,900	,343	1,118	,888 – 1,406			
D3 Delinquent sociaal netwerk	,154	,156	,966	,326	1,166	,858 – 1,585			
D4 Beperkte vrijetijdsbesteding	,096	,134	,511	,475	1,100	,847 – 1,430			
D5 Problematische (ex-) partnerrelatie	-,035	,107	,105	,746	,966	,783 – 1,192			
D6 Instabiliteit woonsituatie	-,043	,144	,088	,767	,958	,722 – 1,272			
D7 Problematisch middelengebruik	,188	,177	1,130	,288	1,207	,853 – 1,708			
D8 Gebrekkige impulsbeheersing	,087	,187	,213	,644	1,090	,755 – 1,574			
D9 Disfunctionele oplossingsvaardigheden	,025	,224	,013	,910	1,026	,661 – 1,590			
D10 Antisociale houding	,046	,155	,089	,766	1,047	,773 – 1,418			
D11 Regelovertredend gedrag	-,210	,196	1,149	,284	,811	,553 – 1,190			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 18. Resultaten Cox regressieanalyse voor de groep in Nederland geboren $n = 976$

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							145,275	17	,000
S1 Leeftijd eerste politiecontact	,111	,068	2,695	,101	1,118	,979 – 1,277			
S2 Aantal eerdere en huidige veroordelingen	,246	,077	10,266	,001	1,279	1,100 – 1,487			
S3 Diversiteit (dreigend) delictgedrag	,024	,097	,060	,807	1,024	,847 – 1,238			
S4 Regelovertredend gedrag in het verleden	-,051	,085	,359	,549	,950	,804 – 1,123			
S5 Instabiliteit opleiding/werk in het verleden	-,015	,085	,032	,858	,985	,834 – 1,163			
S6 Problematisch middelengebruik in het verleden	-,119	,081	1,778	,182	,887	,745 – 1,058			
D1 Disfunctioneren opleiding/werk	,122	,081	2,260	,133	1,130	,964 – 1,325			
D2 Financieel wanbeleid	,001	,073	,000	,989	1,001	,867 – 1,156			
D3 Delinquent sociaal netwerk	,223	,097	5,252	,022	1,250	1,033 – 1,512			
D4 Beperkte vrijetijdsbesteding	,046	,068	,447	,504	1,047	,915 – 1,198			
D5 Problematische (ex-) partnerrelatie	,022	,060	,137	,711	1,022	,910 – 1,149			
D6 Instabiliteit woonsituatie	,275	,074	13,988	,000	1,317	1,140 – 1,512			
D7 Problematisch middelengebruik	,073	,090	,663	,415	1,076	,902 – 1,284			
D8 Gebrekkige impulsbeheersing	-,011	,096	,013	,911	,989	,819 – 1,195			
D9 Disfunctionele oplossingsvaardigheden	-,017	,117	,021	,884	,983	,782 – 1,237			
D10 Antisociale houding	-,091	,106	,748	,387	,913	,742 – 1,123			
D11 Regelovertredend gedrag	,231	,074	9,506	,002	1,260	1,088 – 1,460			

* Standaard fout (b); Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 19. Resultaten Cox regressieanalyse voor de groep mannen $n = 1106$

	B	SE (b)*	Wald	P	Hazard ratio e^b	95% BI van Hazard ratio	χ^2	df	p
							155,487	17	,000
S1 Leeftijd eerste politiecontact	,126	,061	4,327	,038	1,135	1,007 - 1,278			
S2 Aantal eerdere en huidige veroordelingen	,221	,074	8,993	,003	1,248	1,080 - 1,442			
S3 Diversiteit (dreigend) delictgedrag	,045	,090	,247	,619	1,046	,876 - 1,248			
S4 Regelovertredend gedrag in het verleden	,027	,076	,130	,719	1,028	,886 - 1,192			
S5 Instabiliteit opleiding/werk in het verleden	,028	,078	,132	,716	1,029	,883 - 1,199			
S6 Problematisch middelengebruik in het verleden	-,149	,085	3,072	,080	,861	,729 - 1,018			
D1 Disfunctioneren opleiding/werk	,072	,064	,933	,334	1,075	,929 - 1,244			
D2 Financieel wanbeleid	,052	,054	,666	,415	1,053	,930 - 1,193			
D3 Delinquent sociaal netwerk	,179	,068	4,446	,035	1,196	1,013 - 1,412			
D4 Beperkte vrijetijdsbesteding	,040	,084	,398	,528	1,041	,918 - 1,181			
D5 Problematische (ex-) partnerrelatie	,023	,087	,185	,667	1,023	,921 - 1,137			
D6 Instabiliteit woonsituatie	,181	,108	7,166	,007	1,198	1,050 - 1,368			
D7 Problematisch middelengebruik	,127	,092	2,318	,128	1,136	,964 - 1,338			
D8 Gebrekkige impulsbeheersing	-,006	,073	,004	,949	,994	,838 - 1,180			
D9 Disfunctionele oplossingsvaardigheden	-,058	,108	,291	,589	,943	,763 - 1,166			
D10 Antisociale houding	-,043	,092	,221	,638	,958	,800 - 1,146			
D11 Regelovertredend gedrag	,144	,073	3,893	,048	1,155	1,001 - 1,332			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 20. Resultaten Cox regressieanalyse voor de groep vrouwen n = 162

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							27,698	17	,05
S1 Leeftijd eerste politiecontact	-,127	,198	,412	,521	,881	,597 - 1,299			
S2 Aantal eerdere en huidige veroordelingen	,671	,201	11,205	,001	1,957	1,321 - 2,898			
S3 Diversiteit (dreigend) delictgedrag	-,224	,317	,497	,481	,800	,429 - 1,489			
S4 Regelovertredend gedrag in het verleden	-,696	,323	4,640	,031	,498	,264 - ,939			
S5 Instabiliteit opleiding/werk in het verleden	-,070	,277	,063	,802	,933	,542 - 1,606			
S6 Problematisch middelengebruik in het verleden	,282	,239	1,399	,237	1,326	,831 - 2,118			
D1 Disfunctioneren opleiding/werk	-,066	,257	,065	,798	,936	,566 - 1,550			
D2 Financieel wanbeleid	-,109	,214	,260	,610	,896	,589 - 1,365			
D3 Delinquent sociaal netwerk	,152	,284	,286	,593	1,164	,667 - 2,033			
D4 Beperkte vrijetijdsbesteding	,234	,170	1,881	,170	1,263	,905 - 1,764			
D5 Problematische (ex-) partnerrelatie	-,100	,177	,319	,572	,905	,640 - 1,280			
D6 Instabiliteit woonsituatie	,005	,318	,000	,988	1,005	,539 - 1,874			
D7 Problematisch middelengebruik	,091	,293	,096	,757	1,095	,617 - 1,943			
D8 Gebrekkige impulsbeheersing	,548	,295	3,446	,063	1,731	,970 - 3,088			
D9 Disfunctionele oplossingsvaardigheden	-,030	,320	,009	,924	,970	,518 - 1,816			
D10 Antisociale houding	-,111	,300	,136	,713	,895	,497 - 1,612			
D11 Regelovertredend gedrag	,186	,241	,595	,441	1,204	,751 - 1,930			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 21. Resultaten Cox regressieanalyse voor de groep Agressie & Vermogen n = 526

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							56,151	17	,000
S1 Leeftijd eerste politiecontact	-,013	,078	,026	,873	,988	,847 – 1,151			
S2 Aantal eerdere en huidige veroordelingen	,309	,095	10,571	,001	1,363	1,131 – 1,642			
S3 Diversiteit (dreigend) delictgedrag	,050	,119	,174	,676	1,051	,832 – 1,328			
S4 Regelovertredend gedrag in het verleden	-,037	,106	,122	,727	,964	,783 – 1,186			
S5 Instabiliteit opleiding/werk in het verleden	-,013	,101	,017	,896	,987	,809 – 1,204			
S6 Problematisch middelengebruik in het verleden	-,106	,103	1,078	,299	,899	,735 – 1,099			
D1 Disfunctioneren opleiding/werk	,022	,095	,055	,815	1,022	,849 – 1,231			
D2 Financieel wanbeleid	,148	,087	2,854	,091	1,159	,977 – 1,376			
D3 Delinquent sociaal netwerk	,166	,114	2,101	,147	1,180	,943 – 1,476			
D4 Beperkte vrijetijdsbesteding	,029	,089	,105	,746	1,029	,864 – 1,226			
D5 Problematische (ex-) partnerrelatie	-,045	,083	,289	,591	,956	,812 – 1,126			
D6 Instabiliteit woonsituatie	,208	,091	5,269	,022	1,232	1,031 – 1,472			
D7 Problematisch middelengebruik	,006	,105	,004	,952	1,006	,820 – 1,235			
D8 Gebrekkige impulsbeheersing	,012	,120	,009	,923	1,012	,800 – 1,280			
D9 Disfunctionele oplossingsvaardigheden	,053	,150	,122	,727	1,054	,785 – 1,415			
D10 Antisociale houding	-,077	,123	,396	,529	,926	,728 – 1,177			
D11 Regelovertredend gedrag	,126	,098	1,635	,201	1,134	,935 – 1,375			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 22. Resultaten Cox regressieanalyse voor de groep in Huiselijk geweld n = 469

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							69,394	17	,000
S1 Leeftijd eerste politiecontact	,123	,101	1,466	,226	1,131	,927 – 1,379			
S2 Aantal eerdere en huidige veroordelingen	,295	,134	4,815	,028	1,343	1,032 – 1,747			
S3 Diversiteit (dreigend) delictgedrag	-,005	,167	,001	,976	,995	,717 – 1,381			
S4 Regelovertredend gedrag in het verleden	,089	,142	,393	,531	1,093	,827 – 1,445			
S5 Instabiliteit opleiding/werk in het verleden	-,104	,137	,578	,447	,901	,689 – 1,179			
S6 Problematisch middelengebruik in het verleden	-,030	,192	,025	,875	,970	,666 – 1,414			
D1 Disfunctioneren opleiding/werk	,169	,142	1,423	,233	1,184	,897 – 1,563			
D2 Financieel wanbeleid	-,154	,105	2,132	,144	,857	,697 – 1,054			
D3 Delinquent sociaal netwerk	-,144	,186	,599	,439	,866	,601 – 1,247			
D4 Beperkte vrijetijdsbesteding	,071	,106	,443	,506	1,073	,872 – 1,312			
D5 Problematische (ex-) partnerrelatie	-,109	,094	1,325	,250	,897	,745 – 1,079			
D6 Instabiliteit woonsituatie	,251	,128	3,831	,050	1,286	1,000 – 1,654			
D7 Problematisch middelengebruik	,366	,178	4,238	,040	1,442	1,018 – 2,044			
D8 Gebrekkige impulsbeheersing	,130	,188	,477	,490	1,139	,787 – 1,647			
D9 Disfunctionele oplossingsvaardigheden	-,004	,178	,001	,980	,996	,703 – 1,411			
D10 Antisociale houding	,101	,165	,375	,540	1,106	,801 – 1,527			
D11 Regelovertredend gedrag	-,105	,161	,424	,515	,900	,657 – 1,235			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 23. Resultaten Cox regressieanalyse voor de groep Jongvolwassenen (JOVO) n = 142

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	p
							69,980	17	,000
S1 Leeftijd eerste politiecontact	,254	,231	1,208	,272	1,289	,820 – 2,028			
S2 Aantal eerdere en huidige veroordelingen	,363	,225	2,606	,106	1,438	,925 – 2,236			
S3 Diversiteit (dreigend) delictgedrag	,009	,253	,001	,972	1,009	,615 – 1,655			
S4 Regelovertredend gedrag in het verleden	-,286	,220	1,700	,192	,751	,488 – 1,155			
S5 Instabiliteit opleiding/werk in het verleden	,537	,262	4,202	,040	1,711	1,024 – 2,861			
S6 Problematisch middelengebruik in het verleden	-,336	,223	2,274	,132	,714	,461 – 1,106			
D1 Disfunctioneren opleiding/werk	,205	,215	,908	,341	1,228	,805 – 1,873			
D2 Financieel wanbeleid	-,066	,168	,156	,693	,936	,674 – 1,300			
D3 Delinquent sociaal netwerk	,833	,232	12,896	,000	2,301	1,460 – 3,626			
D4 Beperkte vrijetijdsbesteding	,082	,167	,241	,623	1,085	,783 – 1,504			
D5 Problematische (ex-) partnerrelatie	,182	,155	1,371	,242	1,200	,885 – 1,627			
D6 Instabiliteit woonsituatie	-,037	,180	,043	,835	,963	,677 – 1,370			
D7 Problematisch middelengebruik	-,258	,260	,989	,320	,772	,464 – 1,285			
D8 Gebrekkige impulsbeheersing	,119	,180	,435	,510	1,126	,791 – 1,603			
D9 Disfunctionele oplossingsvaardigheden	-,439	,292	2,255	,133	,645	,364 – 1,143			
D10 Antisociale houding	-,312	,246	1,605	,205	,732	,452 – 1,186			
D11 Regelovertredend gedrag	,291	,182	2,545	,111	1,338	,936 – 1,913			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

Tabel 24. Resultaten Cox regressieanalyse voor de groep in behandeling vanwege Seksueel Grensoverschrijdend Gedrag (SGG) N = 133

	B	SE (b)*	Wald	p	Hazard ratio e ^b	95% BI van Hazard ratio	χ^2	df	P
							29,784	17	,024
S1 Leeftijd eerste politiecontact	,757	,537	1,987	,159	2,131	,744 - 6,102			
S2 Aantal eerdere en huidige veroordelingen	-3,164	1,823	3,011	,083	,042	,001 - 1,506			
S3 Diversiteit (dreigend) delictgedrag	3,720	2,203	2,851	,091	41,246	,550 - 3094,2			
S4 Regelovertredend gedrag in het verleden	-,182	1,153	,025	,874	,833	,087 - 7,982			
S5 Instabiliteit opleiding/werk in het verleden	,801	,623	1,652	,199	2,228	,657 - 7,560			
S6 Problematisch middelengebruik in het verleden	1,242	,638	3,789	,052	3,462	,991 - 12,091			
D1 Disfunctioneren opleiding/werk	-,125	,846	,022	,882	,882	,168 - 4,630			
D2 Financieel wanbeleid	,181	,667	,074	,786	1,199	,324 - 4,431			
D3 Delinquent sociaal netwerk	1,123	1,112	1,020	,313	3,074	,348 - 27,694			
D4 Beperkte vrijetijdsbesteding	-1,359	,929	2,138	,144	,257	,042 - 1,588			
D5 Problematische (ex-) partnerrelatie	1,594	,881	3,275	,070	4,925	,876 - 27,694			
D6 Instabiliteit woonsituatie	1,877	1,340	1,962	,161	6,537	,473 - 90,416			
D7 Problematisch middelengebruik	-1,142	,881	1,679	,195	,319	,057 - 1,703			
D8 Gebrekkige impulsbeheersing	-1,478	1,026	2,076	,150	,228	,031 - 1,703			
D9 Disfunctionele oplossingsvaardigheden	-,847	,897	,891	,345	,429	,074 - 2,489			
D10 Antisociale houding	-,561	,931	,363	,547	,570	,092 - 3,539			
D11 Regelovertredend gedrag	,416	1,270	,107	,743	1,515	,126 - 18,250			

* Standaard fout (b), Bonferroni gecorrigeerde $p \leq ,0024$; significante voorspellers na de Bonferroni-correctie zijn dikgedrukt.

8. Actuariële inschatting van het recidiverisico

In de FARE versie 1 is een actuariële risicoclassificatie opgenomen voor de statische FARE-items (S1 – S6). Deze classificatie is gebaseerd op onderzoeksgegevens naar vergelijkbare statische items van de RAF GGZ volwassenen. In dit hoofdstuk worden drie actuariële risicoclassificaties: op basis van de statische risicofactoren, op basis van de dynamische risicofactoren en op basis van alle FARE-items, toegevoegd. Daarvoor zijn drie *Chi-square Automatic Interaction Detection* (CHAID) analyses uitgevoerd op de FARE-metingen die verzameld zijn voor het onderzoek naar de predictieve validiteit.

Onderzoeksgroep

De onderzoeksgroep is dezelfde als die voor de predictieve validiteit werd gebruikt ($N = 1268$).

Analyses

Een reeks CHAID-analyses zijn uitgevoerd om te onderzoeken of een voorspellend model of beslisboom kan worden ontwikkeld om te bepalen hoe de opgenomen risicofactoren in het model het beste samenvloeien om de uitkomst van recidive te voorspellen. Dit gebeurt aan de hand van een beslisboom die volgens een statistisch algoritme tot stand komt. De afhankelijke variabele is recidive nee/ja met score van respectievelijk 0 of 1. De onafhankelijke variabele is de somscore van de scores op de FARE-items, die kan variëren tussen 0 en 68. Op basis van statistische algoritmes worden zogenaamde FARE-afkappunten bepaald in verhouding tot de recidive waardoor subgroepen ontstaan met elk een andere recidivekans. Met andere woorden, er wordt een verhouding berekend tussen variërende somscores op de FARE en de bijhorende recidivepercentages. Dus een relatief lage somscore op de FARE moet in verhouding staan tot het recidivepercentage, dat dus ook relatief laag moet zijn. Hierbij wordt op basis van de grootste chikwadraat-waarde bepaald welke risicofactor (of zoals in dit onderzoek de totaalscore op de FARE) de meeste variantie verklaart.

Resultaten

Tabel 25 beschrijft voor de totale onderzoeksgroep de actuariële risicoinschatting door de scores op alle statische én dynamische risicofactoren bij elkaar op te tellen, hetgeen een totaalscore oplevert die kan variëren tussen 0 en 68 punten. Uit de CHAID-analyse resulteren vijf groepen variërend van zeer laag tot zeer hoog risico op algemene recidive, $\chi^2(4) 119,245$, $p < .001$. De voorspellende waarde van dit model heeft een $AUC = ,71$ ($p < ,001$; BI $,68 - ,75$).

Tabel 25. Actuariële risico inschatting op basis van alle items

Totaalscore op 17 FARE items	Percentage gerecidiveerd	
≤ 14 punten	5,6 %	Zeër laag
15 – 20	11,5%	Laag
21 – 32	20,0%	Matig
33 – 37	30,5%	Hoog
38 – 68	42,9%	Zeër hoog

NB. De minimale score die een cliënt kan behalen op deze risicofactoren is 0 punten en de maximale score is 68.

De beschikbare informatie om een risicotaxatie te kunnen doen, is afhankelijk van de (forensische) behandel- of begeleidingsetting. Soms is er relatief veel informatie over het verleden beschikbaar om de statische risicofactoren te kunnen scoren en soms is deze informatie minimaal. Wanneer er enkel informatie beschikbaar is over de afgelopen zes maanden, wil je als behandelaar of begeleider wel een actuariële risico-inschatting kunnen doen op basis van de dynamische risicofactoren. Deze dynamische risicofactoren kunnen immers gescoord worden op basis van gesprekken met de cliënt en zijn minder afhankelijk van dossierinformatie. Om hierin te voorzien, is een CHAID-analyse uitgevoerd op de statische en dynamische FARE risicofactoren afzonderlijk. De resultaten hiervan zijn weergegeven in tabel 26 voor de statische risicofactoren ($\chi^2(4) 101,185$, $p < .001$) en in tabel 27 voor de dynamische risicofactoren ($\chi^2(3) 77,909$, $p < .001$). De voorspellende waarde voor deze twee modellen zijn respectievelijk $AUC = ,71$ ($p < ,001$; BI $,67 - ,74$) en $AUC = ,68$ ($p < 001$; BI $,64 - ,72$).

Tabel 26. Actuariële risico inschatting op basis van statische items

Totaalscore op 6 statische FARE items	Percentage gerecidiveerd	
≤ 3 punten	3,4 %	Zeer laag
4 – 5	9,1%	Laag
6 – 8	15,1%	Matig
9 – 12	24,3%	Hoog
13 – 24	33,5%	Zeer hoog

NB. De minimale score die een cliënt kan behalen op deze statische risicofactoren is 0 punten en de maximale score is 24.

Dynamische risicofactoren

Op basis van de dynamische risicofactoren werden vier significant verschillende groepen onderscheiden.

Tabel 27. Actuariële risico inschatting op basis van dynamische items

Totaalscore op 11 dynamische FARE items	Percentage gerecidiveerd	
≤ 9 punten	6,7 %	Zeer laag
10 – 15	13,4 %	Laag
16 – 23	24,3 %	Matig
24 – 44	36,2 %	Hoog

NB. De minimale score die een cliënt kan behalen op deze dynamische risicofactoren is 0 punten en de maximale score is 44. Op basis van de dynamische risicofactoren werden enkel vier significant verschillende groepen onderscheiden.

9. Evaluatie bruikbaarheid van de FARE

9.1 Methode

Dataverzameling Evaluatie gebruik FARE

De evaluatie is aan alle behandelaren van Transfore ($n = 75$) en de Waag ($n = 300$) aangeboden. Bij de andere instellingen, Fivoor Ambulant, Kairos en Verslavingszorg Noord Nederland (VNN) was de implementatie van de FARE nog niet voltooid. De behandelaren werden gevraagd om de verschillende onderdelen van de FARE, zoals de beschrijving van de indicatoren en de scorerichtlijnen op basis van vijf stellingen over het gebruiksgemak en helderheid van de tekst met een cijfer te beoordelen en toe te lichten.

Begin 2018, enkele weken na de implementatie van de FARE, werd aan professionals gevraagd naar hun ervaringen met de FARE in de dagelijkse praktijk. De resultaten van deze evaluatie zijn gebruikt voor de verbetering van de FARE-handleiding versie 1 en de digitale versie in de ROM-systemen.

In totaal hebben 87 behandelaren het evaluatieformulier ingevuld (23,2% van de benaderde behandelaren), 19 (22,9%) behandelaren van Transfore en 64 (77,1%) van de Waag. Vier reacties zijn niet meegenomen, twee behandelaren hadden de FARE niet bij volwassen cliënten ingevuld en één behandelaar had de evaluatie twee keer ingevuld. De twee behandelaren die de FARE hadden afgenomen bij minderjarige cliënten vonden de FARE een prettig instrument, maar enkel bruikbaar als er factoren opgenomen zouden worden specifiek voor de jeugdpopulatie. Hun voorkeur ging uit naar een kort instrument vergelijkbaar met de FARE maar dan voor jeugd, die ook gebruikt kan worden om de behandelvoortgang te kunnen blijven volgen als jongeren de overgang naar volwassenheid maken. De behandelaren bij Transfore hadden met vooral 1-5 metingen minder ervaring met de FARE dan de behandelaren van de Waag met overwegend meer dan 10 metingen (zie Tabel 28).

Tabel 28. Aantal keer FARE ingevuld

Aantal taxaties	Totaal		Transfore		Waag	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1-5	23	27,7	12	63,2	11	17,2
6-9	13	15,7	5	26,3	8	12,5
Meer dan 10	47	56,6	2	10,5	45	70,3
Totaal	83	100	19	100	64	100

9.2 Resultaten

Het gebruiksgemak van de FARE werd beoordeeld met een 7,5 (SD 1,1; range 3-10), ook de indicatoren, items en antwoordcategorieën werden met een ruim voldoende beoordeeld. Men vond het moeilijker om behandeldoelen te formuleren op basis van de FARE-items, een kwart gaf hieraan een onvoldoende (zie Tabel 29).

Tabel 29. Cijfers per stelling van de FARE

	Gem. Cijfer	≤5	6	7	8	9	10	Geen cijfer
1. Het gebruik van de FARE is over het algemeen gemakkelijk	7,5 (SD 1,1; 3-10)	4,8%	9,6%	21,7%	53,0%	7,2%	2,4%	1,3%
2. De items van de FARE sluiten aan bij de situatie van cliënten in de ambulante setting	7,1 (SD 1,4; 2-9)	9,6%	14,5%	24,1%	43,4%	4,8%	-	3,6%
3. De indicatoren van de FARE zijn helder en compleet omschreven	7,2 (SD 1,2; 3-10)	6,0%	18,1%	24,1%	44,6%	3,6%	1,2%	2,4%
4. De antwoordcategorieën van de FARE zijn helder omschreven	7,3 (SD 1,3; 3-10)	7,2%	14,5%	30,1%	34,9%	10,8%	2,4%	-
5. Het is gemakkelijk om met de items van de FARE behandeldoelen te formuleren	6,1 (SD 1,5; 1-8)	27,7%	24,1%	27,7%	14,5%	-	-	6%

Verbeterpunten

Er zijn diverse verbeterpunten genoemd, bij zowel de onvoldoendes als voor extra onderwerpen zoals training en ROM-systemen.

De formulering van items, zoals Opleiding / werk in het heden en verleden, (S5 en D1), Sociaal netwerk (D3), Vrijtijdsbesteding (D4), Gebrekkige impulsbeheersing (D8) en Disfunctionele oplossingsvaardigheden (D9) moesten duidelijker worden geformuleerd zodat ze niet voor interpretatie vatbaar zouden zijn. Men verwachtte ook een instructie hoe de items gescoord moesten worden als cliënten in detentie zitten en de mogelijkheid om een item niet te scoren vanwege onvoldoende informatie. Kritiek op de scorerichtlijnen van items varieerde van te specifiek geformuleerd tot te breed geformuleerd. Het meest werd aangegeven dat de situatie van de cliënt in geen enkele score optie letterlijk past. Bijvoorbeeld een cliënt had nooit problemen op zijn werk maar werd wel ontslagen na één maar zeer heftig conflict. Een conflict wordt niet als problematische gezien (score 1) maar met ontslag scoor je al snel een '4', beide scores beschrijven niet de gehele situatie.

Behandelaren misten de volgende specifieke risicofactoren en domeinen: domein over psychische en persoonlijkheidsproblematiek, gezinssituatie, opvoeding en gezin van herkomst, zeker bij partnergeweld en kindermishandeling, niet geregistreerde delicten en de ernst van de delicten. Behandelaren misten ook specifieke risicofactoren met betrekking tot seksueel grensoverschrijdend gedrag.

Behandeldoelen

In vergelijking met de andere vragen is een grote groep ontevreden (27,7% geeft een onvoldoende) over de relatie tussen de FARE en behandeldoelen. Voor een deel gaat de onvrede over het ontbreken van een specifiek invulveld voor behandeldoelen (n = 4). Een andere groep acht de FARE minder geschikt voor het formuleren van specifieke behandeldoelen voor doelgroepen zoals huiselijk geweldplegers, jeugdige cliënten en zedendelinquenten (n = 3). Dit valt samen met het gerapporteerde gemis van risicofactoren specifiek voor deze doelgroepen. Tot slot is er een groep die risicotaxatie-uitkomsten in het algemeen (ongeacht het instrument) niet (altijd) geschikt of nuttig achten voor het formuleren van behandeldoelen (n = 3).

ROM-systemen

Twee behandelaren gaven expliciet aan dat zij het verwarrend vonden dat de cijfers 0 - 4 niet voor de tekstuele scorerichtlijnen in de ROM-systemen zoals QuestManager en User staan omdat deze wel voor de scorerichtlijnen in de handleiding staan. Er is behoefte aan een invulveld voor de behandeldoelen en een maatwerkrapportage met grafieken. In QuestManager staat de knop 'volgende vraag' onder de knop 'definitief', hierdoor drukken behandelaren soms op de verkeerde knop, waardoor de FARE niet meer te bewerken is. Verder is er behoefte aan automatisch gegenereerde behandeldoelen of instructie wanneer een behandeldoel geformuleerd moet worden vanaf score 'x' (signaalfunctie of vlagfunctie).

Training

Alle behandelaren (die deelnamen aan de evaluatie) gaven aan ervaring te hebben in het verrichten van risicotaxaties met andere instrumenten dan de FARE. Toch hebben deze behandelaren behoefte aan een FARE-training. Vooral behandelaren met een LS/CMI achtergrond misten een degelijke training zoals ze gewend zijn bij de LS/CMI.

Uit een deel van de kritiekpunten blijkt ook dat enkele aan het onderzoek deelnemende behandelaren (n = 3) de handleiding niet helemaal gelezen hebben. Bijvoorbeeld hoe de optelling werkt in de actuariële scoring en hoe de actuariële scoring zich verhoudt tot de klinische inschatting (includeren van de dynamische risicofactoren). Tot slot gaven behandelaren (n = 2) te kennen dat de FARE versie 1 te sterk gericht is op de statische factoren.

10. FARE versie 2

De FARE Versie 2 (FARE V2; Van Horn, Eisenberg, Bouman, Van den Hanenberg, & Bogaerts, 2020) is een aanpassing van de eerste versie van de FARE (Van Horn et al., 2016c). De aanpassingen zijn gebaseerd op uitkomsten de gebruikersevaluatie, de onderzoeksresultaten van de interbeoordelaarsbetrouwbaarheid van de FARE versie 1 en de formulering van de actuariële risicoclassificatie op basis van alle FARE-items. Items zijn tekstueel aangescherpt om de indicatoren en de scorerichtlijnen te verduidelijken, de actuariële inschatting van het recidiverisico is aangescherpt op basis van recente data en er is een uitgebreide instructie geschreven voor het formuleren en onderbouwen van het recidiverisico. Een gedetailleerde beschrijving van de tekstuele wijzigingen zijn opgenomen in Bijlage 2.

Formeel behoort dit deelonderzoek niet tot het door KFZ geaccordeerde onderzoek. De onderzoekers zien echter een meerwaarde in het informeren van de gebruikers over de psychometrische kwaliteiten (minimaal de interbeoordelaarsbetrouwbaarheid) van de aangepaste versie van de FARE. In dit onderzoek is gekeken of de aangepaste FARE versie 2 een verbetering heeft opgeleverd van de interbeoordelaarsbetrouwbaarheid.

10.2 Methode

Onderzoeksgroep

Van de 84 geïnterviewde cliënten van de onderzoeksgroep ten behoeve van het onderzoek naar de interbeoordelaarsbetrouwbaarheid van de FARE versie 1 (Hoofdstuk 4) zijn 30 cliënten random geselecteerd voor het onderzoek naar de interbeoordelaarsbetrouwbaarheid van de FARE versie 2.

Beoordelaars en procedure

De FARE's versie 2 zijn gescoord door een onderzoeker, de tweede auteur van dit rapport en een Masterstudent Forensische psychologie die vooraf getraind werd in het gebruik van de FARE. De beoordelaars hebben onafhankelijk van elkaar de interviewopnames beluisterd en de FARE versie 2 gescoord. De ervaring met de FARE van beide beoordelaars was ondanks hun functie vergelijkbaar. De student was niet betrokken bij de dataverzameling van de interbeoordelaarsbetrouwbaarheid van FARE versie 1.

Power

De berekening van de noodzakelijke grootte van de onderzoeksgroep is opnieuw uitgevoerd. De verwachte betrouwbaarheid is aangepast aan de hypothese dat de verbeteringen aan het instrument zullen leiden tot een hogere overeenstemming. Bij een minimale acceptabele betrouwbaarheid (ρ_0) vastgesteld op $ICC = ,50$ en de verwachte betrouwbaarheid (ρ_1) $ICC = ,80$, twee beoordelaren per cliënt, een alpha van $,05$ (one-tailed) en een Power (β) van $,80$ (Arifin, 2020; zie ook Bujang & Baharum, 2017) is het aantal benodigde cliënten $N = 28$.

Analyses

De overeenstemming tussen beoordelaars werd berekend met Intraclass Correlation Coefficients, two-way random model en absolute agreement (ICC's, Koo & Li, 2016). De 'single measures' zijn gerapporteerd omdat de FARE uiteindelijk door een beoordelaar gescoord gaat worden. De ICC is berekend op item niveau (de 6 statische en 11 dynamische items) en op risico-inschatting niveau (actuarieel en klinisch gestructureerd). Voor de interpretatie van de resultaten zijn de richtlijnen van Koo en Li (2016) gebruikt, waarbij een $ICC < ,50$ slecht is; $,50 \leq ICC < ,75$ matig; $,75 \leq ICC < ,90$ is goed en een $ICC \geq ,90$ is uitstekend.

Resultaten

De interbeoordelaarsbetrouwbaarheid van de FARE versie 2 is met ICC's tussen de $,56$ en $,96$ (matig – uitstekend) over het geheel verbeterd ten opzichte van de interbeoordelaarsbetrouwbaarheid van de FARE versie 1. Voor de FARE versie 1 gold voor beide beoordelaarskoppels dat de ICC voor drie dynamische risicofactoren slecht was (zie D4, D9 en D11 in Tabel 28). En tussen behandelaren en studenten was ook de ICC voor D6 Instabiliteit woonsituatie en de klinische inschatting van het recidive risico slecht. Deze factoren zijn alle verbeterd naar een matige tot goede overeenstemming (zie Tabel 30).

Tabel 30. Intraclass Correlation Coëfficiënten (ICC) voor de items van de FARE versie 2 (N = 30) vergeleken met de ICC voor de items van FARE versie 1 (N = 42)

	FARE Versie 1			FARE Versie 2		
	ICC	,95 BI	n	ICC	,95 BI	n
Statische risicofactoren						
S1 Leeftijd eerste politiecontact	,942***	,895 - ,968	42	,935***	,868 - ,969	29
S2 Aantal eerdere en huidige veroordelingen	,947***	,904 - ,971	42	,933***	,864 - ,967	30
S3 Diversiteit (dreigend) delictgedrag	,700***	,507 - ,826	42	,805***	,634 - ,902	30
S4 Regelovertredend gedrag verleden	,751***	,585 - ,857	42	,824***	,658 - ,913	29
S5 Instabiliteit opleiding/werk verleden	,737***	,564 - ,848	42	,796***	,615 - ,898	30
S6 Problematisch middelengebruik verleden	,822***	,694 - ,900	42	,960***	,917 - ,981	29
Dynamische risicofactoren						
D1 Disfunctioneren opleiding/werk	,709***	,518 - ,832	42	,670***	,416 - ,828	30
D2 Financieel wanbeleid	,711***	,521 - ,834	41	,737***	,515 - ,866	30
D3 Delinquent sociaal netwerk	,614***	,387 - ,771	42	,927***	,852 - ,964	30
D4 Beperkte vrijetijdsbesteding	,411***	,125 - ,634	41	,690***	,447 - ,839	30
D5 Problematische (ex-) partnerrelatie	,859***	,750 - ,922	41	,811***	,585 - ,913	29
D6 Instabiliteit woonsituatie	,625***	,399 - ,780	41	,711***	,469 - ,855	28
D7 Problematisch middelengebruik	,539***	,278 - ,725	41	,865***	,721 - ,936	29
D8 Gebrekkige impulsbeheersing	,612***	,384 - ,770	42	,614***	,286 - ,804	30
D9 Disfunctionele oplossingsvaardigheden	,461**	,182 - ,670	42	,560***	,264 - ,762	30
D10 Antisociale houding	,654***	,436 - ,798	42	,687***	,374 - ,848	30
D11 Regelovertredend gedrag	,148 ^{ns} **	-,148 - ,421	42	,647***	,379 - ,815	30
Inschatting recidiverisico						
Actuariële inschatting recidiverisico	,860***	,755 - ,922	42	,936***	,868 - ,969	27
Klinische inschatting recidiverisico	,626***	,392 - ,784	40	,687***	,436 - ,838	30

BI = betrouwbaarheidsinterval ns = niet significant

11. Conclusie

Met dit onderzoek zijn enkele psychometrische eigenschappen van de FARE onderzocht en is de bruikbaarheid van het instrument in kaart gebracht. Dit onderzoek richtte zich op de beantwoording van de volgende zes onderzoeksvragen:

1. Wat is de overeenstemming tussen beoordelaars op de FARE (interbeoordelaarsbetrouwbaarheid)?
2. Wat is de samenhang van de FARE met andere risicotaxatie-instrumenten die hetzelfde beogen te meten (convergente validiteit)?
3. In welke mate wordt in herhaalmetingen met de FARE verandering bij cliënten gemeten?
4. In hoeverre zijn de risicofactoren van de FARE afzonderlijk en de FARE als geheel voorspellend voor recidive?
5. Zijn er subgroepen en/of cliëntprofielen te onderscheiden en zijn er verschillen in predictieve validiteit tussen subgroepen?
6. Is er aanleiding om op basis van de data verzameld met de FARE de actuariële risico inschatting bij te stellen?

In de volgende onderdelen worden per onderzoeksvraag de resultaten kort samengevat.

Interbeoordelaarsbetrouwbaarheid (onderzoeksvraag 1)

Het deelonderzoek naar de interbeoordelaarsbetrouwbaarheid van de FARE versie 1 richtte zich op de vraag in welke mate twee beoordelaars op basis van dezelfde informatie onafhankelijk van elkaar tot dezelfde scores komen. De beoordelaarskoppels bestonden uit studenten en behandelaars enerzijds en studenten onderling anderzijds. De resultaten laten zien dat bij alle beoordelaarskoppels de overeenstemming op de statische items beter (veelal matig tot zeer goede overeenstemming) is dan op de dynamische items (veelal slechte tot matige overeenstemming). De betere overeenstemming op de statische risicofactoren kan deels worden verklaard door de strakkere scoringsrichtlijnen van de statische risicofactoren. De scoring op het item Leeftijd eerste politiecontact bijvoorbeeld vraagt om de leeftijd waarop client voor het eerst in aanraking is geweest met de politie als verdachte. Dergelijke items laten weinig ruimte voor eigen interpretatie. Hetzelfde geldt voor een aantal andere statische FARE-items. Met betrekking tot de dynamische FARE items zijn de scoringsrichtlijnen over het algemeen minder eenduidig opgesteld, waardoor meer ruimte is voor interpretatie. Voor een ander deel kan de slechtere overeenstemming ook te maken hebben met tekstuele onduidelijkheden in de indicatoren en de scoringsrichtlijnen. Een voorbeeld hiervan is het item Regelovertredend gedrag waar zowel de mate van houden aan regels als de mate van motivatie voor behandeling moet worden gescoord. Deze twee onderdelen zijn niet noodzakelijkerwijs in gelijke mate problematisch voor een cliënt, waardoor het scoren bemoeilijkt wordt. In versie 2 van de FARE zijn deze tekstuele onduidelijkheden in diverse dynamische items verholpen, hetgeen de interbeoordelaarsbetrouwbaarheid ten goede zou moeten komen.

Convergente validiteit (onderzoeksvraag 2)

De convergente validiteit gemeten als de correlatie tussen de FARE en het IFpBE, LS/CMI en RAF GGZ was over het algemeen sterk, behalve voor zes – voornamelijk dynamische – FARE-items. Een inhoudelijke verkenning doet vermoeden dat de FARE-items en de inhoudelijk meest overeenkomende items uit de andere taxatie-instrumenten toch niet hetzelfde beogen te meten. Zo ligt bijvoorbeeld bij het IFpBE item 5 (draagt de patiënt zorg voor een evenwichtige dagindeling?) de focus op dagbesteding (hoe vol plant iemand zijn tijd?) en bij de FARE item D4 (beperkte vrijetijdsbesteding) ligt de nadruk meer op pro-sociale besteding van de vrije tijd en op inbedding in de maatschappij. Geconcludeerd wordt dat het bepalen van de convergente validiteit van een risicotaxatie-instrument met andere risicotaxatie-instrumenten niet zonder beperkingen is. Een belangrijke beperking is dat voor het meten van recidiverisico geen 'gouden standaard' bestaat. Daarom werd er voor gekozen om meerdere andere risicotaxatie-instrumenten in de analyse te betrekken en zo zorgvuldig mogelijk de items inhoudelijk te matchen. Dat is lastiger gebleken dan verwacht. De convergente validiteit zou mogelijk beter onderzocht kunnen worden door zelfrapportage vragenlijsten die inhoudelijk dezelfde concepten meten te correleren aan het risico-item. Een gebrekkige impulsbeheersing bijvoorbeeld kan gerelateerd worden aan een gevalideerde zelfrapportage vragenlijst waarin het concept impulsiviteit wordt gemeten, zoals de Barrett Impulsivity Scale (BIS; Patton, Stanford, Barratt, 1995), of aan response inhibitie taken, zoals de Iowa Gambling taak (IGT; Bechara, Damasio, Damasio, & Anderson, 1994).

Veranderbaarheid (onderzoeksvraag 3)

De veranderbaarheid van de dynamische FARE-items is onderzocht aan de hand van herhaalmetingen bij 2180 cliënten. Voor de gehele onderzoeksgroep zijn de veranderingen op de dynamische risicofactoren en de gestructureerde klinische inschatting significant, maar met een klein effect. De grootste verbetering is zichtbaar bij de risicofactoren problematische (ex-)partnerrelatie, disfunctionele oplossingsvaardigheden en de gestructureerde klinische inschatting van het recidiverisico met een veranderscore van ,03 per maand. Voor de gehele onderzoeksgroep geldt dat op de langere termijn de verbetering afneemt, waarna er een kleine verslechtering optreedt die vervolgens weer afbuigt richting een verbetering (een S-vormige verandering).

De startscores voor de totale onderzoeksgroep zijn relatief laag, waardoor logischerwijs de relatieve verandering per maand ook klein is. Anders gezegd, cliënten die zonder al te hoge scores op de dynamische risicofactoren aan hun behandeling beginnen, zullen relatief weinig veranderingen laten zien in de loop van de tijd. Omdat de relatief kleine verandering gebaseerd is op de totale groep (laagscoorders en hoogscoorders middelen elkaar uit, hetgeen zich uit in een matige score) hebben we de analyse opnieuw gedaan maar dan voor enkel de hoogscoorders (score 2 of hoger). De veronderstelling was dat er grotere veranderingen zouden worden gevonden bij deze groep die reeds aan het begin van de behandeling tot de meer problematische groep behoorden. De resultaten wezen inderdaad uit dat de hoogscoorders op alle risicofactoren verbeterden en dat de veranderingen groter waren dan bij de gehele onderzoeksgroep. De grootste verandering is zichtbaar voor de risicofactor problematische (ex-) partnerrelatie (veranderscore van 0,08 per maand).

Over het algemeen kan worden geconcludeerd dat de dynamische risicofactoren gevoelig zijn voor verandering en dat de veranderingen sterker zijn bij de meest problematische groep. De sterkste veranderingen bij zowel de totale groep als de hoogscoorders vonden plaats in de eerste maanden van de behandeling. Om te onderzoeken of de aangetoonde verandering statistisch relevant was, werden de Reliable Change Indices (RCIs) berekend voor twee groepen: een groep die de behandeling afmaakt en een subgroep hiervan die een hoge score had. De resultaten wezen dat een cliënt na afronding van de behandeling op de FARE-items een statistisch betrouwbare verandering doormaakt als de problematiek met twee tot drie punten afneemt. Als specifiek gekeken wordt naar de groep cliënten met een hoge startscore blijkt dat er gesproken kan worden van een statische betrouwbare verbetering als de score op een item met een tot twee punten afneemt.

In deze deelstudie is de interne sensitiviteit van de FARE onderzocht. Met andere woorden, onderzocht werd of de FARE items gevoelig zijn voor het vaststellen van veranderingen. Of verandering op de een of meerdere FARE-items ook een verandering in recidive bewerkstelligt (ook wel externe sensitiviteit genoemd, Viljoen, Shaffer, Gray, & Douglas, 2017) moet in toekomstig onderzoek worden vastgesteld.

Predictieve validiteit (onderzoeksvraag 4 en 5)

De predictieve validiteit van de FARE is op twee manieren (ROC-analyses en Cox regressieanalyses) onderzocht voor de gehele onderzoeksgroep en voor acht aparte subgroepen onderzocht. Met ROC-analyses is onderzocht of met de FARE accuraat voorspeld kan worden of cliënten zullen recidiveren. Daarnaast is met de Cox regressieanalyse onderzocht of de FARE-items en de inschatting van het recidiverisico niet alleen voorspelt of iemand zal recidiveren maar ook of de snelheid waarmee cliënten recidiveren hierbij een rol speelt.

Voor de gehele onderzoeksgroep was het recidivepercentage 17,1%. Dit is een relatief laag percentage vergeleken met andere forensische populaties waarbij percentages gevonden werden van 19% bij tbs-gestelden, 45% bij ex-gedetineerden en 80% bij ISD-ers (Drieschner, Hill & Weijters, 2018). Het recidivepercentage van 17,1% betreft enkel nieuwe rechtbankzaken en is mogelijk een onderschatting omdat niet elk (recidive-)delict leidt tot een politiecontact en/of veroordeling (Scurich & John, 2019). Bijkomend is dat zeer lage recidivepercentages de kans op vals negatieve resultaten vergroten, dat wil zeggen dat op basis van de resultaten onterecht wordt aangenomen dat een FARE-item geen of nauwelijks een voorspellende waarde heeft. .

Uit zowel de ROC-analyses als de Cox regressieanalyses bleek dat een betere voorspellende waarde bereikt wordt als de scores van de risicofactoren worden opgeteld (somscore) of als alle risicofactoren tezamen worden opgenomen in het voorspellingsmodel. De ROC-analyse wees verder uit dat in de meeste subgroepen de risicofactoren onafhankelijk van elkaar recidive slecht tot matig voorspellen, maar dat de risicofactor Eerdere en huidige veroordelingen wel een significante

bijdrage levert aan de voorspelling van recidive(-snelheid). Voor mannen, vrouwen en cliënten in behandeling vanwege seksueel grensoverschrijdend gedrag worden geen factoren met een individuele bijdrage aan de voorspelling van recidivesnelheid gevonden. Op basis van de bovengenoemde resultaten kan worden gesteld dat de FARE als geheel een betere voorspelling biedt voor algemene recidive dan de afzonderlijke risicofactoren. Demografische en delictgerelateerde informatie zoals geboorteland, geslacht of type delict voegen hier geen voorspellingskracht aan toe.

Tot slot, de resultaten moeten worden gezien vanuit de beperking dat recidive alleen onderzocht kon worden voor de populatie van de Waag. Omdat de FARE breed wordt gebruikt in het ambulante forensische werkveld is het van belang om de voorspellende waarde in een grotere populatie met meer instellingen te onderzoeken waarbij een langere follow-up tijd wordt aangehouden.

Actuariële risicoclassificatie (onderzoeksvraag 6)

Met een *Chi-square Automatic Interaction Detection (CHAID)* analyse is onderzocht of het mogelijk is om met behulp van een optelsom van de scores op alle FARE-items recidiverisicogroepen te onderscheiden. Er zijn drie verschillende actuariële risicoclassificaties geformuleerd: 1) op alle risicofactoren van de FARE, 2) op enkel de statische risicofactoren en 3) op enkel de dynamische factoren. De drie typen risicoclassificaties leveren een goed hulpmiddel op voor het indelen van de populatie in risicogroepen en deze zijn opgenomen in de FARE versie 2.

Het recidiverisico in de onderzochte groep is relatief laag, 17,1%. Dit is een vergelijkbaar percentage dat eerder werd gevonden in een grotere Waaggroep (N = 5204) met een langere follow-up tijd (19,3% binnen 2 jaar en 27,9% binnen 6,6 jaar; Eisenberg, et al., 2015). Omdat de CHAID-analyse gebaseerd is op dit recidivepercentage zie je dat in de groep met de hoogste score op de FARE (zeer hoog recidiverisico) 'maar' 42,9% gerecidiveerd. Echter, als over het algemeen het recidivepercentage al laag ligt, is 42,9% kans op recidive nog altijd relatief hoog vergeleken met de andere groepen.

Ervaring met gebruik van de FARE

Om de ervaringen van behandelaren met de FARE in de dagelijkse praktijk te onderzoeken is een evaluatie uitgevoerd bij alle behandelaren van De Waag en Transfore, welke is ingevuld door 87 behandelaren. Deze evaluatie is enkele weken na de implementatie van de FARE in 2017 verstuurd. Het gebruiksgemak van de FARE werd met een 7,5 beoordeeld, ook de beschrijving van de indicatoren, scorerichtlijnen en aansluiting van de FARE bij de ambulante forensische praktijk werd beoordeeld met een 7,1.

Een viertal verbeterpunten werd genoemd: (1) de items, disfunctioneren opleiding/werk (S5 en D1), sociaal netwerk (D3), beperkte vrijetijdsbesteding (D4), gebrekkige impulsbeheersing (D8) en disfunctionele oplossingsvaardigheden (D9) moesten duidelijker omschreven worden zodat ze niet voor interpretatie vatbaar zijn, (2) de behandelaren misten een aantal items/domeinen (bijvoorbeeld specifieke risicofactoren bij huiselijk geweld, zeden en persoonlijkheidsproblematiek), (3). Behandelaren gaven aan dat de scoringsrichtlijnen niet voor elke cliënt passend zijn, (4) behandelaren misten de vertaalslag van risicofactoren naar behandeldoelen. Verschillende van deze aspecten zijn verwerkt in versie 2 van de FARE: enkele van de genoemde items zijn, ook vanwege de lage interbeoordelaarsbetrouwbaarheid aangescherpt. Zo is bijvoorbeeld in de beschrijving van de scoringsrichtlijnen aangegeven dat het scoren van de situatie van een cliënt aan de hand van de opgenomen indicatoren dient te geschieden, en dat het zelden zo zal zijn dat er een perfecte match is. De beoordelaar wordt geadviseerd de cliënt in de "geest" van het item te beoordelen, en in de toelichting de overwegingen op te nemen. Ook is het advies in geval van twijfel het onderwerp expliciet te agenderen voor intercollegiale toetsing. Voor wat betreft de vertaling van de itemscores naar behandeldoelen is in de versie 2 van de FARE een formulier opgenomen ter ondersteuning van dit proces.

FARE versie 2

De resultaten van de interbeoordelaarsbetrouwbaarheid, de actuariële risicoclassificatie en de suggesties ter verbetering uit het onderzoek naar de bruikbaarheid van de FARE, zijn gebruikt om de indicatoren en scorerichtlijnen van de risicofactoren aan te scherpen.

Vervolgens is de interbeoordelaarsbetrouwbaarheid opnieuw onderzocht. Deze is over het geheel verbeterd met matige tot uitstekende overeenstemming voor de items en de gestructureerde klinische inschatting van het recidiverisico. Deze mate van overeenstemming is vergelijkbaar met eerdere studies met de HCR, HKT-30, het IFpBE, de LC/CMI en de RAF GGZ (Bogaerts, et al.,

2018; Canton, et al., 2004; Hildebrand, et al., 2005; Rocque & Plubber-Beale, 2014; Schuringa, et al., 2014; Van Horn, et al., 2009). Het onderzoek naar de interbeoordelaarsbetrouwbaarheid van versie 2 is uitgevoerd door een onderzoeker en een student. Deze studie moet opnieuw uitgevoerd worden met behandelaren als beoordelaars om vast te stellen of ook zij op een verbeterde overeenstemming uitkomen.

Advies bij het gebruik van de FARE

De FARE is een generiek instrument wat inhoudt dat het algemene recidive voorspelt. De FARE is sinds januari 2019 opgenomen in de lijst van risicotaxatie-instrumenten die als veldnorm gelden voor het ambulante forensische veld (Prestatie Indicatorengids 2019). In de evaluatie van de FARE kwam echter naar voren dat de FARE onvoldoende geschikt is voor bepaalde typen plegers, zoals zedendelinquenten en huiselijk geweldplegers. Behandelaren wordt geadviseerd voor specifieke doelgroepen naast de FARE doelgroep-specifieke risicotaxatie instrumenten af te nemen zoals de Static-Stable-Acute of de B-Safer.

Daarnaast kan in het gebruik van de FARE de interbeoordelaarsbetrouwbaarheid verhoogd worden door het volgen van een FARE-training en door periodiek afstemming te zoeken met collega's, bijvoorbeeld in intervisies en multidisciplinaire overleggen.

Voor het formuleren van een adequaat risicomangementplan op cliëntniveau wordt aangeraden om met een (delict-)analyse (zie KFZ call 2019-111 verkorte delictanalyse) de samenhang tussen risicofactoren (onderling) en het delictgedrag in kaart te brengen.

Vervolgstappen

Uit het onderzoek naar de interbeoordelaarsbetrouwbaarheid, de convergente validiteit, de veranderbaarheid en de predictieve validiteit komt enige evidentie naar voren die deze psychometrische kwaliteiten van de FARE voldoende ondersteunen. De FARE is het enige risicotaxatie- en risico-evaluatie instrument in Nederland dat op deze wijze is onderzocht in het ambulante forensische veld. Het is een instrument dat blijvend zal worden verbeterd. Vooralsnog is de voorzichtige conclusie dat de FARE een geschikt instrument is voor risicotaxatie- en ROM in het forensische ambulante veld.

Uit het onderzoek komen verbeterpunten naar voren die kunnen bijdragen aan de verbetering van de onderzochte psychometrische eigenschappen en de gebruiksvriendelijkheid van het instrument. Deze verbeterpunten zijn verwerkt in de FARE versie 2. Daarnaast wordt de FARE spin ontwikkeld, een webapplicatie van de dynamische FARE-items die zowel door de behandelaar als de cliënt ingevuld kan worden. De scores op de FARE-items worden visueel weergegeven in de vorm van een 'spinnenweb'. De cliënt kan de uitkomsten delen met de behandelaren en zo kunnen zij in gesprek gaan over risico- en beschermende factoren.

De vertaling van de uitkomsten van de risicotaxatie naar behandeldoelen is een ingewikkeld klinisch proces die buiten de scope van onderhavig project valt. Daarom is gestart met een inventarisatie van de mogelijkheden voor een vertaalslag van de risicotaxatie naar risicomangement voor het gehele forensische veld (KFZ-project 2019-100, Bouman e.a., 2020). Vervolgonderzoek, door andere onderzoekers dan de auteurs, zou gericht moeten worden op het opnieuw onderzoeken van de psychometrische kwaliteiten en praktische bruikbaarheid van de FARE versie 2 bij de verschillende type zorgverleners binnen het bredere ambulante forensische veld en aanpalende sectoren, waaronder de Forensische FACT, RIBW's, Maatschappelijke opvang. Daarnaast zal de predictieve validiteit van de FARE versie 2 binnen verschillende ambulante forensische poliklinieken moeten worden onderzocht.

12. Referenties

- Ægisdóttir, S., Spengler, P. M., & White, M. J. (2006) Should I Pack My Umbrella? Clinical Versus Statistical Prediction of Mental Health Decisions. *The Counseling Psychologist, Vol 34*, No. 3, 410 – 419. doi: 10.1177/0011000006286696
- Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Wormith, J. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency, 52*, 7-27.
- Arifin, W. N. (2020). Sample size calculator (web). Retrieved from <http://wnarifin.github.io>
- Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*, 7-15.
- Blok, G., De Beurs, E., De Ranitz, A., & Rinne, T. (2010). Psychometrische stand van zaken van risicotaxatie-instrumenten voor volwassenen in Nederland. *Tijdschrift voor Psychiatrie, 52*(5), 331-341.
- Bogaerts, S., Spreen, M., Ter Horst, P, & Gerlsma, C. (2018). Predictive validity of the HKT-R Risk Assessment for two and five-year recidivism in a cohort of Dutch Forensic Psychiatric Patients. *International Journal of offender Therapy and Comparative Criminology, 62*, 2259-2270.
- Bonta, J. L., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior, 19*, 278-287. doi:10.1016/j.avb.2014.04.014
- Bonta, J. & Andrews, D. A. (2017). *The psychology of criminal conduct (6th ed.)*. New York, NY: Routledge.
- Bouman, Y.H.A., Schene, A.H., & Ruiter, C. de (2008). Subjective well-being and recidivism in forensic psychiatric outpatients: Preliminary evidence for a good lives model. *International Journal of Forensic Mental Health, 8*, 225 - 234. doi: 10.1080/14999011003635647
- Bouman, Y. H. A., Wismeyer, C. E., Lammers, S. M. M., & Philipse, M. W. G. (2012). Level of Service/Case Management Inventory (LS/CMI). Nederlandse geautoriseerde (onderzoeks)versie. Nijmegen: Kairos.
- Bujang, & Baharum, (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Sciences, 12*(1), 1-11.
- Campbell, M. A., French, S., & Gendreau, P. (2007). *Assessing the utility of risk assessment tools and personality measures in the prediction of violent recidivism for adults*. Department of Public Safety and Emergency Preparedness: Ottawa, Canada.
- Canton, W. J., Van der Veer, T., Van Panhuis, V., Verheul, R., & Van der Brink, W. (2004). De betrouwbaarheid van risicotaxatie in de pro Justitia rapportage. Een onderzoek met behulp van de HKT-30. *Tijdschrift voor Psychiatrie, 46*(8), 537-542.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- De Vogel, V., & de Ruiter, C. (2003). Verschillen tussen onderzoekers en behandelaars in het inschatten van het risico van gewelddadig gedrag. *DITH 23*, 22. Doi:10.1007/BF03060302
- De Vogel, V., & De Ruiter, C. (2005). The HCR-20 in personality disordered female offenders: a comparison with a matched sample of males. *Clinical Psychology and Psychiatry*(12), 226-240.
- De Vogel, V., De Ruiter, C., Hildebrand, M., Bos, B., & Van de Ven, P. (2004). Type of Discharge and Risk of Recidivism Measured by the HCR-20: A Retrospective Study in a Dutch Sample of Treated Forensic Psychiatric Patients. *International Journal of Forensic Mental Health*(3), 149-165.
- Divisie Forensische Zorg en Justitiële Jeugdinstellingen (2019). Kernset prestatie-indicatoren Forensische Psychiatrie Forensische geestelijke gezondheidszorg, verslavingszorg en zorg voor mensen met een (licht) verstandelijke beperking. Den Haag.
- Douglas, K.S., James, R. P., Ogloff, J. D., & Hart, S. D. (2003). Evaluation of a Model of Violence Risk Assessment Among Forensic Psychiatric Patients. *Psychiatric Services, 54*(10), 1372 – 1379. doi:10.1176/appi.ps.54.10.1372
- Eisenberg, M. J., Van Horn, J.E., Van der Put, C. E., Hendriks, J., & Stams, G.J.J.M (2015). Vooronderzoek 'Kernset Risicofactoren Ambulant behandelde delinquenten. Voorspellers van recidive bij ambulant behandelde zeden-, (huiselijk)gewelds- en vermogensdelinquenten. Den Haag: KFZ.

- Eisenberg, M. J., Van Horn, J.E., Dekker, J. M., Assink, M., Van der Put, C. E., Hendriks, J., & Stams, G. J. J. M. (2019). Static and Dynamic Predictors of General and Violent Criminal Offense Recidivism in the Forensic Outpatient Population. A Meta-Analysis. *Criminal Justice and Behavior*, *Vol. 46*: 5, 732 – 750. doi: 10.1177/0093854819826109
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral and biomedical sciences. (39), 175-191.
- Grieger, L., & Hosser, D. (2014). Which risk factors are really predictive? An analysis of Andrews and Bonta's "Central Eight" risk factors for recidivism in German youth correctional facility inmates. *Criminal Justice and Behavior*, *41*, 613-634.
- Gutierrez, L., Wilson, H. A., Rugge, T., & Bonta, J. (2013). The prediction of recidivism with aboriginal offenders: A theoretically informed meta-analysis. *Canadian Journal of Criminology and Criminal Justice*, *55*, 55-99. doi:10.3138/cjccj.2011.E.51
- Guy, L. S., Douglas, K. S., & Hart, S. D. (2015). Risk assessment and communication. In B. L. Cutler & P. A. Zapf (Eds.), *APA handbooks in psychology®. APA handbook of forensic psychology, Vol. 1. Individual and situational influences in criminal and civil contexts* (p. 35–86). American Psychological Association. doi: 10.1037/14461-003
- Harte, J., & Breukink, M. (2010). Objectiviteit of schijnzekerheid? Kwaliteit, mogelijkheden en beperkingen van instrumenten voor risicotaxatie. *Tijdschrift voor Criminologie*, *52*(1), 52-72.
- Heck, R. H., Thomas, S. L., & Tabata, L. N. (2014). *Multilevel and Longitudinal Modeling with IBM SPSS. Quantative methodology series. (2nd ed.)*. London: Routledge. doi: 10.4324/9780203701249
- Hildebrand, M., Hesper, B. L., Spreen, M., & Nijman, H. (2005). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie. Een onderzoek naar de betrouwbaarheid en predictieve validiteit van de HCR-20, HKT-30 en PCL-R*. Utrecht: Expertisecentrum Forensische Psychiatrie.
- Holden, J. E., Kelley, K., & Agarwal, R. (2008). Analyzing Change: A Primer on Multilevel Models with Applications to Nephrology. *American Journal of Nephrology*(28), 792-801.
- Jung, S., Ledi, D., & Daniels, M. K. (2013). Evaluating the concurrent validity of the HCR-20 scales. *Journal of Risk Research*, *16*(6), 697-711, doi: 10.1080/13669877.2012.726247
- Koo & Li, (2016). A guideline of selecting and reporting Intraclass Correlation Coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155-63. doi: 10.1016/j.jcm.2016.02.012
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. doi:10.2307/2529310
- Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Nijman, H., De Kruijk, C., & Van Nieuwenhuizen, C. (2004). Behavioral changes during forensic psychiatric (TBS) treatment in the Netherlands. *International Journal of Law and Psychiatry*, *27*(1), 79-85. *International Journal of Law and Psychiatry*(27), 79-85.
- Nikulin, M. & Wu, H.-D., I., (2016). *The Cox Model and its applications*. Berlin: Springer. doi: 10.1007/978-3-662-49332-8
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*, 156-176. doi:10.1037/a0035080
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology*, *51*:768-774.
- 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*, 615–620. doi:10.1007/s10979-005-6832-7
- Rocque, M., & Plummer-Beale, J. (2014). In the eye of the beholder? An examination of the inter-rater reliability of the LSI-R and YLS/CMI in a correctional agency. *Journal of Criminal Justice*, *42*(6), 568-578.
- Scurich, N., & John, R. (2019). The dark figure of sexual recidivism. *Behavioral Sciences & the Law*, *37*(2), 158–175. <https://doi.org/10.1002/bsl.2400>
- Schönberger, H., Hildebrand, M., Spreen, M., & Bloem, O. (2008). *De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie bij seksuele delinquenten*. Utrecht: Expertisecentrum Forensische Psychiatrie.
- Schuringa, E. (2011). Een voorbeeldcasus uit het Patiënt Volg Systeem van het FPC Dr. S. van Mesdag. *GGZet Wetenschappelijk*, *15*(2), 55-69.

- Schuringa, E., Heininga, V. E., Spreen, M., & Bogaerts, S. (2018). Concurrent and predictive validity of the instrument for forensic treatment evaluation: From risk assessment to routine, multidisciplinary treatment evaluation. *International Journal of Offender Therapy and Comparative Criminology*, *62*, 1281-1299. doi: 10.1177/0306624X16676100
- Schuringa, E., Spreen, M., & Bogaerts, S. (2014). Inter-rater and test-retest reliability, internal consistency and factorial structure of the Dutch Instrument for Forensic Treatment Evaluation. *Journal of Forensic Psychology Practice*, *14*, 127-144.
- Spreen, M., Brand, E., Ter Horst, P., & Bogaerts, S. (2014). *Handleiding en Methodologische Verantwoording HKT-R, Historische, Klinische en Toekomstige - Revisie*. Groningen: Dr. van Mesdag kliniek.
- Sunhee, P., & Lake, E. T. (2005). Multilevel modelling of a clustered continuous outcome: Nurses' work hours and burnout. *Nursing Research* (54).
- Troquete, N. A. C., Brink, R. H. S. van den, Beintema, H., Mulder, T., Os, T. W. D. P. van, Schoevers, R. A., & Wiersma, D. (2015). Predictive Validity of the Short-Term Assessment of Risk and Treatability for Violent Behavior in Outpatient Forensic Psychiatric Patients. *Psychological Assessment*, Vol. 27, No. 2, 377 – 391.
- Van den Brink, R.H.S., Troquete, N.A.C., Beintema, H. et al. Risk assessment by client and case manager for shared decision making in outpatient forensic psychiatry. *BMC Psychiatry* **15**, 120 (2015). <https://doi.org/10.1186/s12888-015-0500-3>
- Van der Knaap, L., Leenarts, L., & Nijssen, L. (2007). *Psychometrische kwaliteiten van de Recidive Inschattingsschalen (RISc): Interbeoordelaarsbetrouwbaarheid, interne consistentie en congruente validiteit*. Den Haag: WODC.
- Van der Veecken, F. C. A., Lucieer, J., & Bogaerts, S. (2016). Routine outcome monitoring and clinical decision-making in Forensic Psychiatry based on the Instrument for Forensic Treatment Evaluation. *Plos one*.
- Van der Veecken, F. C. A., Lucieer, J., & Bogaerts, S. (2018). Forensic psychiatric treatment evaluation: The clinical evaluation of treatment progress with repeated forensic routine outcome monitoring measures. *International Journal of Law and Psychiatry*, *57*, 9-16. doi: 10.1016/j.ijlp.2017.12.002
- Van Horn, J., Bogaerts, S., Eisenberg, M., Van der Put, C., Dekker, J., Van den Hanenberg, F., & Bouman, Y. (2016a). Kernset K-factoren voor het ambulante forensische veld: Een multi-method onderzoek naar risico- en beschermende factoren in relatie tot algemene recidive, geweldsrecidive en seksuele recidive. *Kwaliteit Forensische Zorg: Utrecht*.
- Van Horn, J., Eisenberg, M., Bouman, Y., Van den Hanenberg, F., Van der Put, C., & Bogaerts, S. (2016b). Eindverslag pilot 'praktische bruikbaarheid Forensisch Ambulante Risico Evaluatie - FARE. *Kwaliteit Forensische Zorg: Utrecht*.
- Van Horn, J., Eisenberg, M., Bouman, Y., Van den Hanenberg, F., & Bogaerts, S. (2016c). Handleiding Forensisch Ambulante Risico Evaluatie - FARE. *Kwaliteit Forensische Zorg: Utrecht*.
- Van Horn, J., Eisenberg, M., & Uzieblo, K. (2016). Risicotaxatie in de Nederlandse ambulante forensische geestelijke gezondheidszorg. *Tijdschrift voor Psychiatrie*, *58*(8), 583-592.
- Van Horn, J., Wilpert, J., Bos, M., Eisenberg, M., & Mulder, J. (2009). WaagSchaal Jeugd – De psychometrische kwaliteit van een gestructureerd klinisch risicotaxatie-instrument voor de ambulante forensische psychiatrie. *Panopticon. Tijdschrift voor Strafrecht, criminologie en forensisch welzijnswerk*, *30*(2), 23-34.
- Van Horn, J., Wilpert, J., Scholing, A., & Mulder, J. (2008). *Waagschaal (thans RAF GGZ) volwassenen. Risicotaxatie instrument voor de Ambulante Forensische Psychiatrie. Handleiding. Versie 2008*. De Waag: Utrecht.
- Van Horn, J., Eisenberg, M., Bouman, Y., Van den Hanenberg, F., & Bogaerts, S. (2020). Handleiding Forensisch Ambulante Risico Evaluatie versie 2 – FARE versie 2. *Kwaliteit Forensische Zorg: Utrecht*.
- Viljoen, J., Shaffer, C., Gray, A., & Douglas, K. (2017). Are adolescent risk assessment tools sensitive to change? A framework and examination of the SAVRY and the YLS/CMI. *Law and Human Behavior*, *41*(3), 244–257. <https://doi.org/10.1037/lhb0000238>
- Wilson, H. A., & Gutierrez, L. (2014). Does one size fit all? A meta-analysis examining the predictive ability of the Level of Service Inventory (LSI) with Aboriginal offenders. *Criminal Justice and Behavior*, *41*, 196-219. doi:10.1177/0093854813500958

- Wooditch, A., Tang, L. L., & Taxman, F. S. (2014). Which criminogenic need changes are most important in promoting desistance from crime and substance use? *Criminal Justice and Behavior, 41*, 276-299. doi:10.1177/0093854813503543
- Wormith, J. S., Hogg, S., & Guzzo, L. (2012). The predictive validity of a general risk/needs assessment inventory on sexual offender recidivism and an exploration of the professional override. *Criminal Justice and Behavior, Vol. 39*, No. 12, 1511 – 1538.

Bijlage 1. Overeenkomende items FARE, IFpBE, LS/CMI en RAF GGZ

FARE-items met overeenkomende IFpBE items	
S1 t/m S6	Geen overeenkomende IFpBE items
D1. Disfunctioneren opleiding/werk	6. Toont de patiënt goede arbeidsvaardigheden?
D2. Financieel wanbeleid	9. Toont de patiënt voldoende financiële vaardigheden? 10. Toont de patiënt voldoende administratieve vaardigheden?
D3. Delinquent sociaal netwerk	17. Is de patiënt georiënteerd op personen of subculturen met antisociale en/of criminele opvattingen?
D4. Beperkte vrijetijdsbesteding	5. Draagt de patiënt zorg voor een evenwichtige dagindeling?
D5. Problematische (ex-)partnerrelatie	20. Heeft de patiënt een acceptabele partnerrelatie?
D6. Instabiliteit woonsituatie	Geen overeenkomende IFpBE items
D7. Problematisch middelengebruik	M1. Erkent de patiënt het middelengebruik als een probleem? M3. Toont de patiënt vaardigheden om middelengebruik te voorkomen? M4. Is er afgelopen periode sprake geweest van middelengebruik?
D8. Gebrekkige impulsbeheersing	11. Vertoont de patiënt impulsief gedrag?
D9. Disfunctionele oplossingsvaardigheden	4. Toont de patiënt over het algemeen adequate coping vaardigheden?
D10. Antisociale houding	3. Erkent de patiënt zijn delict en neemt hij hier verantwoordelijkheid voor? 12. Vertoont de patiënt vijandigheid en op welke wijze? 13. Vertoont de patiënt asociaal gedrag ten aanzien van de hulpverlener (en de organisatie)?
D11. Regelovertredend gedrag	2. Werkt de patiënt mee aan jouw behandeling? 16. Houdt de patiënt zich aan de regels en voorwaarden?
Actuariële inschatting recidiverisico	Geen inschatting recidiverisico in het IFpBE
Klinische inschatting recidiverisico	

FARE-items met overeenkomende LS/CMI items	
S1. Leeftijd eerste politiecontact	5. Arrestaties voor 16 jaar?
S2. Aantal eerdere en huidige veroordelingen	1. Eerdere veroordelingen gehad?
S3. Diversiteit delictgedrag	Geen overeenkomende LS/CMI items
S4. Regelovertredend gedrag in het verleden	8. Aanklacht of schending reclasseringsvoorwaarden?
S5. Instabiliteit opleiding/werk verleden	10. Vaak werkloos? 11. Nooit een vol jaar aan het werk geweest? 12. ≤ 3 jaar voorgezet onderwijs? 14. Ooit geschorst of van school gestuurd?
S6. Problematisch middelengebruik in het verleden	28. Ooit een alcoholprobleem gehad 29. Ooit een drugsprobleem gehad
D1. Disfunctioneren opleiding/werk	9. Op dit moment < 20u betaald werk per week? 15. Deelname/prestatie 16. Omgang met medeleerlingen/collegae 17. Omgang met leraren/superieuren
D2. Financieel wanbeleid	43a. Financiële problemen

D3. Delinquent sociaal netwerk	21. Strafolblad in de naaste familie? 24. Enige bekenden met crimineel gedrag 25. Enige vrienden met crimineel gedrag 26. Heeft < 3 vrienden zonder crimineel gedrag 27. Heeft <3 bekenden zonder crimineel gedrag
D4. Beperkte vrijetijdsbesteding	22. Geen recente deelname aan clubs/organisaties 23. Mate van actieve vrijetijdsbesteding
D5. Problematische (ex)partnerrelatie	18. Mate van tevredenheid over de huidige leefvorm
D6. Instabiliteit woonsituatie	Geen overeenkomende LS/CMI items.
D7. Problematisch middelengebruik	30. Mate van actuele drankproblemen 31. Mate van actuele drugsproblemen 32. Relatie drank/drugs met criminaliteit? 33. Familie klaag over drank/drugsgebruik 34. Problemen op school/werk door drank/drugsgebruik 35. Medische problemen als gevolg van drank/drugsgebruik
D8. Gebrekkige impulsbeheersing	Geen overeenkomende LS/CMI items
D9. Disfunctionele oplossingsvaardigheden	Geen overeenkomende LS/CMI items
D10. Antisociale houding	36. Goedkeuring van delict(en) 37. Afwijzing van maatschappelijke conventies 40. (vermoeden van) antisociale trekken
D11. Regelovertredend gedrag	Geen overeenkomende LS/CMI items
Actuariële inschatting recidiverisico	Risico/behoefte niveau op basis van totaalscore Sectie 1
Klinische inschatting recidiverisico	Geen klinische inschatting recidiverisico in de LS/CMI

FARE-items met overeenkomende RAF GGZ items	
S1. Leeftijd eerste politiecontact	D1f. Jonge leeftijd eerste uiting antisociaal gedrag
S2. Aantal eerdere en huidige veroordelingen	D1a. Veroordelingen/ afdoeningen
S3. Diversiteit delictgedrag	D1a. Veroordelingen/ afdoeningen D1b. Officiële aangifte(n)/registraties D1c. Niet off. gevestigd. delicten /wangedragingen
S4. Regelovertredend gedrag in het verleden	D1h. Schending van voorwaarden D11a. Hulpverlening in het verleden
S5. Instabiliteit opleiding/werk verleden	D2a. Afgeronde opleiding voortgezet onderwijs D2b. Ooit geschorst of van school gestuurd D2f. Problemen m.b.t. het arbeidsverleden
S6. Problematisch middelengebruik in het verleden	D8a. Drankmisbruik/-afhankelijkheid verleden D8b. Drugs misbruik/-afhankelijkheid verleden
D1. Disfunctioneren opleiding/werk	D2d. Interesse in/binding met school D2e. Gedragsproblemen op school D2g. Op dit moment werkloos D2h. Werkprestatie D2i. Omgang met collegae D2j. Omgang met leidinggevenden
D2. Financieel wanbeleid	D3a. Ontvangt uitkering D3b. Schulden
D3. Delinquent sociaal netwerk	D6b. Vrienden/kennissen met crimineel verleden
D4. Beperkte vrijetijdsbesteding	D7a. Individuele vrijetijdsbesteding D7b. Contextuele vrijetijdsbesteding
D5. Problematische (ex-)partnerrelatie	D5a. Instabiliteit van relaties
D6. Instabiliteit woonsituatie	D4a. Stabiliteit van de woonsituatie D4b. Achterstandsbuurt
D7. Problematisch middelengebruik	D8c. Drankmisbruik/-afhankelijkheid heden D8d. Drugs misbruik/-afhankelijkheid heden D8e. Diagnose middelen D8f. Interpersoonlijke problemen

	D8g. Problemen op school/werk
D8. Gebrekkige impulsbeheersing	D9e. Impulsiviteit
D9. Disfunctionele oplossingsvaardigheden	D9g. Coping vaardigheden
D10. Antisociale houding	D10a. Goedkeuring of vergoelijking delicten D10b. Minimaliseren of ontkenning delicten
D11. Regelovertredend gedrag	Geen overeenkomende RAF GGZ items
Actuariële inschatting recidiverisico	Geen actuariële recidiverisico inschatting in de RAF GGZ
Klinische inschatting recidiverisico	Hoogst ingeschatte recidiverisico van de zes delict categorieën

Bijlage 2 Overzicht wijzigingen FARE V2

Item	Wijziging
S1. Leeftijd eerste politiecontact	Specificatie dat waarschuwingen en aanhoudingen niet meegeteld worden in dit item. Het gaat dus om de leeftijd waarop de eerste arrestatie, in verzekeringstelling of voorlopige hechtenis van cliënt plaatsvond.
S2. Aantal eerdere en huidige veroordelingen	Specificaties dat veroordelingen pas geteld worden als zij onherroepelijk zijn en dat het enkel om strafrechtelijke veroordelingen gaat, niet om kantonzaken of civielrechtelijke zaken.
S3. Diversiteit dreigend delictgedrag	De definitie van 'dreigend delictgedrag' is aangescherpt.
S4. Regelovertredend gedrag verleden	Alleen frequentie wordt meegenomen en ernst is buiten beschouwing gelaten. Omdat het maken van een indeling op basis van zowel ernst als frequentie van regelovertredend gedrag niet inhoudelijk te onderbouwen is. Is eenmalig een ernstige recidive tijdens de proeftijd erger dan meerdere behandelafspraken niet nakomen? Omdat hier geen rangorde in te maken is, is ervoor gekozen om alleen naar de frequentie te kijken. In hoeverre is er sprake van een patroon van regelovertredend gedrag?
S5. en D1. Instabiliteit opleiding/werk (in het verleden)	Er is gekozen om de scheiding tussen opleiding en werk te verwijderen. In principe zijn dit twee gebieden die dezelfde vaardigheden vereisen (bijv. omgang met autoriteit) en waar ook dezelfde soort problemen zich kunnen voordoen. Bijvoorbeeld: spijbelen en werkverzuim vallen allebei onder ongeoorloofde afwezigheid, alleen in een andere setting. De levensfase bepaalt waar cliënt op dat moment mee bezig is. Je beoordeelt dus het functioneren op beide gebieden.
S6. Problematisch middelengebruik in het verleden	Het statische item bleek onvoldoende onderscheidend van het dynamische item 'D7. Problematisch middelengebruik' omdat er te veel overlap was: tel het aantal DSM 5 criteria over de gehele voorgeschiedenis tot nu óf in het afgelopen half jaar. Het dynamische item blijft gelijk maar bij het statische item wordt in versie 2 alleen gescoord in hoeverre de cliënt in een bepaalde periode van zijn/haar leven aan de vier hoofdcriteria van de DSM 5 voldeed. Als een cliënt voldoet aan criterium 3 (risicogebruik) dan voldoet hij/zij volgens de DSM 5 automatisch ook aan criterium 1 en 2 (beperkte controle en sociale beperkingen). Daarnaast stond bij score 2 en 3 de toevoeging 'of gebruik is in langdurige/vroege remissie'. Dit is in versie 2 weggehaald. Het gaat er bij S6 om of cliënt ooit een periode heeft gehad waarin hij/zij voldeed aan een of meerdere van de beschreven criteria. Hoe lang geleden dat heeft plaatsgevonden neem je niet mee bij de scoring. Je scoort dus de hele voorgeschiedenis van cliënt.
D2. Financieel wanbeleid	Bij score 1 is het volgende toegevoegd: 'de financiën van cliënt werden beheerd door een curator / bewindvoerder / mentor en heeft in de afgelopen 6 maanden daarbij goed gefunctioneerd'. Dit klinkt wellicht tegenstrijdig, omdat het feit dat cliënt onder curatele/bewindvoering/mentorschap staat, betekent dat er sprake is geweest van financieel wanbeleid. Echter, het is een dynamisch item en dus wordt het functioneren in de afgelopen 6 maanden gescoord. Als cliënt dus onder deze financiële begeleiding goed functioneerde in de afgelopen 6 maanden, wordt een score 1 gegeven.
D3. Delinquent sociaal netwerk	De focus lag bij versie 1 op de verhouding tussen prosociale en criminele netwerkleden. In versie 2 wordt alleen nog gekeken naar de criminele netwerkleden en in hoeverre zij cliënt beïnvloeden. Of een cliënt dus 3 vrienden heeft die alle drie crimineel zijn, of dat cliënt 10 vrienden heeft waarvan er 3 crimineel zijn: dat maakt voor de scoring niet uit. Het gaat er namelijk om in hoeverre zij cliënt negatief beïnvloeden. Dit betekent ook dat cliënten die sociaal geïsoleerd zijn de score 0 krijgen (want: als je geen netwerkleden hebt, heb je automatisch ook geen criminele netwerkleden). Score 0 impliceert dat dit gebied potentieel beschermend is, maar dat is dan

	dus niet het geval. Sociaal geïsoleerd zijn is geen beschermende factor, maar het is ook geen risicofactor. Om deze reden is score 0 hier gepast.
D4. Beperkte vrijetijdsbesteding	In versie 1 werd naast de diversiteit, de aard (passief of actief) van de activiteiten en of het met anderen was meegenomen. In versie 2 is zowel de aard als de het al dan niet samen met anderen de activiteiten doen buiten beschouwing gelaten. De nadruk ligt nu op: besteedt cliënt zijn/haar vrije tijd aan prosociale activiteiten of aan criminele, antisociale activiteiten? Of iemand televisie kijkt of naar een voetbaltraining gaat maakt geen verschil aangezien het beide niet-crimineel activiteiten zijn.
D6. Instabiliteit woonsituatie	Bij de eerste indicator 'instabiliteit verblijfadres' werden een aantal voorbeelden genoemd van instabiele verblijfadressen, zoals onderhuur of een kraakpand. Dit was echter vooral gericht op de toekomst: kan cliënt hier voorlopig blijven wonen, of dreigt hij/zij te moeten verhuizen? Omdat dynamische items kijken naar de afgelopen 6 maanden tot de huidige taxatiedatum, is ervoor gekozen om in versie 2 alleen nog te focussen op het aantal verhuizingen in de afgelopen 6 maanden. Je kijkt dus niet vooruit, maar terug.
D7. Problematisch middelengebruik	Bij score 2 en 3 stond de toevoeging 'of gebruik is in langdurige remissie'. Deze toevoeging is in versie 2 weggehaald omdat het dynamische item over de afgelopen 6 maanden gaat. Voor de scoring van dit item kijk je dus naar de aanwezigheid van de 11 criteria in de afgelopen 6 maanden. Als de cliënt in de afgelopen 6 maanden in remissie was, zal hij/zij laag scoren op het dynamische item en hoog op het statische item S6.
D8. Gebrekkige impulsbeheersing	De indicator 'impulsieve keuzes' is verwijderd omdat het onderscheid tussen keuzes en gedrag vaak gering is. Daarom is er in versie 2 voor gekozen om de voorbeelden van impulsieve keuzes samen te voegen met de indicator 'onbeheerst gedrag'.
D9. Disfunctionele oplossingsvaardigheden	Bij dit item zijn de antwoord categorieën aangepast. Het aantal levensgebieden wordt geteld waar de disfunctionele oplossingsvaardigheden voor problemen zorgen. Hier is voor gekozen om de ernst/reikwijdte van de problematiek in kaart te kunnen brengen. Op hoe meer levensgebieden de disfunctionele oplossingsvaardigheden voor verstoring zorgen, hoe instabieler je leven is.
D10. Antisociale houding	De scoringsinstructies zijn verduidelijkt door de zin 'Cliënt liet in de afgelopen 6 maanden geen/af en toe/regelmatig/... antisociale houding zien' toe te voegen. Daarna worden de manieren waarop zich dit uit beschreven.
D11. Regelovertredend gedrag	De inhoudelijke wijzigingen bij dit item zijn hetzelfde als bij item S4. Daarnaast zijn de scoringsrichtlijnen gewijzigd. In versie 1 lag de nadruk bij de scoringsrichtlijnen te veel op behandelbaarheid en de houding van cliënt in de behandeling. Bij dit item moet echter ook regelovertredend gedrag op andere gebieden meegenomen worden (in begeleidingstrajecten, voorwaarden en (strafrechtelijke) bepalingen). In versie 2 worden deze gebieden ook benoemd in de scoringsrichtlijnen omschreven.