



Jeugdige verdachten beter in beeld?

Een haalbaarheidsstudie naar de
betrouwbaarheid van de Ritax-data
uit het Landelijk Instrumentarium
Jeugdstrafrechtken voor
vergelijkingen over de tijd

Cahier 2025-12

G.J. Koops-Geuze
A.M. van der Laan

Jeugdige verdachten beter in beeld?

Een haalbaarheidsstudie naar de
betrouwbaarheid van de Ritax-data
uit het Landelijk Instrumentarium
Jeugdstrafrechtkenen voor
vergelijkingen over de tijd

Cahier 2025-12

G.J. Koops-Geuze
A.M. Van der Laan

Cahier

De reeks Cahier omvat de rapporten van onderzoek dat door en in opdracht van het Wetenschappelijk Onderzoek- en Datacentrum is verricht. Opname in de reeks betekent niet dat de inhoud van de rapporten het standpunt van de Minister van Justitie en Veiligheid weergeeft.

Inhoud

Samenvatting	6
1 Inleiding	9
1.1 Achtergrond	10
1.1.1 Beleidscontext	10
1.1.2 Doel- en vraagstelling	11
1.2 Leeswijzer	13
2 Methoden	14
2.1 Afbakeningen	14
2.2 Methode	15
2.2.1 Betrouwbaarheid	15
2.2.2 Data governance	18
3 Resultaten	20
3.1 Literatuurstudie	20
3.1.1 Zoekopdracht en documentselectie	20
3.1.2 De samenstelling van het LIJ-instrument	21
3.1.3 Eerder onderzoek naar aanpassingen en de betrouwbaarheid van het LIJ	22
3.1.4 Ongewijzigde items over de tijd op basis van de grijze literatuur	24
3.1.5 Tussenconclusie	26
3.2 Interviews en focusgroepen — betrouwbaarheidsvraag	28
3.2.1 (In)consistentie in de afname en invulling	28
3.2.2 Veranderingen in de registratie van afnames over de tijd	29
3.2.3 Bruikbaarheid en beperkingen voor vergelijkingen over de tijd	30
3.2.4 Tussenconclusie	31
3.3 Ontbrekende waarden analyse – Betrouwbaarheidsvraag	32
3.3.1 Data verkennen en definiëren van ontbrekende waarden	32
3.3.2 Achtergrondkenmerken	34
3.3.3 Ontbrekende waarden op itemniveau	35
3.3.4 Ontbrekende waarden op itemniveau binnen verschillende subgroepen	40
3.3.5 Tussenconclusie	44
3.4 Interviews en focusgroepen — Data governance-vraag	46
3.4.1 Inhoudelijke noodzaak	46
3.4.2 Tussenconclusie	48
4 Slot	49
4.1 Beantwoording onderzoeksvragen	49
4.2 Mogelijkheden en beperkingen van het onderzoek	51
4.3 Discussie	52
Summary	56
Literatuur	59

Bijlage 1	Aantallen en percentages van ontbrekende waarden op alle items in de dataset van 2023 en 2022	61
Bijlage 2	Alle items met een significant verband voor het jaar 2023 en 2022 op basis van sekse, leeftijd, geslacht en afnameregio	62

Samenvatting

In deze haalbaarheidsstudie is onderzocht in hoeverre Ritax-data van de Raad voor de Kinderbescherming, die onderdeel zijn van het Landelijk Instrumentarium Jeugd (het LIJ), geschikt zijn om ontwikkelingen in kenmerken van jeugdige verdachten over de tijd te onderzoeken. Bijvoorbeeld als aanvullende databron voor de Monitor Jeugdcriminaliteit (MJC), waarin tweejaarlijks trends in jeugdcriminaliteit worden onderzocht. De beschikbare landelijke politieregistraties bieden nauwelijks inzicht in kenmerken van jeugdige verdachten, zoals relevante risico- en beschermende factoren voor delinquent gedrag. Zelfrapportagegegevens zoals gemeten in de Monitor Zelfgerapporteerde Jeugddelinquentie bieden die informatie wel, maar die richt zich op algemene populaties. De Ritax-data zou deze leemte potentieel voor alle jeugdige verdachten kunnen vullen vanwege de screening op risico- en beschermende factoren die een mogelijke rol spelen in het voorkomen van en terugvallen in criminaliteit. Het is echter niet bekend hoe consistent gegevens uit de Ritax landelijk en over de tijd worden geregistreerd, wat voor vergelijkingen over de tijd wel nodig is. Voordat uitspraken gedaan kunnen worden over veranderingen in kenmerken van jeugdige verdachten over de tijd (of verschillen tussen groepen) is dergelijk inzicht wel belangrijk. Daarom richtte dit onderzoek zich op de bruikbaarheid van de Ritax voor vergelijkingen over de tijd op groeps- en populatieniveau bij jeugdige verdachten. Het ging hierbij om de betrouwbaarheid van de Ritax met betrekking tot de wijze van invulling en volledigheid van deze gegevens in de landelijk beschikbare registratiesystemen.

Om de betrouwbaarheid en bruikbaarheid van de Ritax-data te beoordelen, werd een mixed methods-onderzoek uitgevoerd. Als eerste werd een literatuurstudie gedaan naar wat al bekend was uit eerdere onderzoeken over de constructie, betrouwbaarheid en bruikbaarheid van het LIJ in het algemeen en de Ritax in het bijzonder. Vervolgens werd aan de hand van kwalitatieve, expertinterviews en focusgroepen met raadsonderzoekers, beleidsmedewerkers en praktijkprofessionals de wijze van afname, invulling en bruikbaarheid van de Ritax in de praktijk onderzocht. Ten derde werd door middel van kwantitatieve analyses van geanonimiseerde Ritax-datasets gekeken naar de vulling en volledigheid van geregistreerde gegevens op itemniveau, voor de jaren 2022 en 2023. Door deze combinatie van methoden is inzicht verkregen in inhoudelijke en praktische veranderingen aan de Ritax over de tijd. Tevens werden hierdoor patronen van ontbrekende gegevens en variatie in registratie tussen domeinen en subgroepen zichtbaar. Aanvullend is door middel van twee focusgroepen ook een *data governance*-vraag besproken waarbij het specifiek ging over wat mogelijk en nodig zou zijn om Ritax-data op microniveau structureel beschikbaar te hebben bij het WODC, indien deze voldoende bruikbaar zou zijn.

Uit de resultaten van het kwalitatieve gedeelte blijkt dat de consistentie van Ritax-registraties over de tijd beperkt blijkt. Inhoudelijke wijzigingen in het instrument, zoals de invoering van de Ritax 2.0 in 2022, hebben geleid tot een methodebreuk. Hierdoor zijn gegevens van voor deze periode moeilijk vergelijkbaar met de latere gegevens. Data vanaf de Ritax 2.0-periode zou mogelijk wel geschikt zijn om ontwikkelingen over de tijd te onderzoeken. Uit de interviews en focusgroepen bleek echter dat er in de praktijk regionale verschillen zijn in werkwijzen bij de afnames van de Ritax waarbij niet duidelijk is of die verschillen over de tijd consistent zijn. Deze regionale variatie in de manier van registreren en de veranderingen in de Ritax over de tijd beperken de

bruikbaarheid ervan voor onderzoek naar vergelijkingen tussen regio's en subgroepen en veranderingen in kenmerken van jeugdige verdachten over de tijd. Aanvullend bleek uit de subgroepanalyses in het kwantitatieve gedeelte dat de mate van ontbrekende waarden voor sommige items significant verschilde naar geslacht, leeftijd en regio van afname.

Een bevinding op basis van de kwantitatieve analyses was dat de Ritax-data niet altijd volledig zijn ingevuld, gezien de mate van 'onbekend'-scores. Vooral binnen de domeinen gezin, attitude/houding en relaties komen relatief veel 'onbekend'-scores voor, wat vragen oproept over de bruikbaarheid van de risico- en beschermende factoren binnen deze domeinen voor onderzoeksdoeleinden. Hoewel de mate van ontbrekende waarden iets lager was voor 2023 dan voor 2022, is mogelijke selectiviteit in de Ritax-data een aandachtspunt. Verschillen in ontbrekende waarden gekeken naar geslacht, leeftijd en regio betekenen dat de data incompleet zijn. Dat heeft gevolgen voor de representativiteit van de gegevens en beperkt vergelijkingen over de tijd. Het is denkbaar dat de gegevens in het ene jaar of in een bepaalde regio vooral betrekking hebben op een specifieke subpopulatie, terwijl in een ander jaar of een andere regio een andere subpopulatie domineert. Deze mogelijke verschuivingen in de onderliggende populatie bemoeilijken interpretaties van trends. Dergelijke selectiviteit kan alleen worden beoordeeld door Ritax-data op microniveau te koppelen aan een externe bron die de volledige populatie jeugdige verdachten (of jeugdige justitiabelen) omvat.

Verder liet een analyse van demografische kenmerken zien dat leeftijd, geslacht, delict zwaarte en regionale spreiding grotendeels gelijk bleven voor de jaren 2022 en 2023. De meerderheid van de gegevens betreft jongens (ongeveer 90%), met een gemiddelde leeftijd van 15,4 jaar. Omdat slechts twee jaren zijn vergeleken en er sprake is van een aanzienlijk aantal ontbrekende waarden, kan weinig worden geconcludeerd over veranderingen in delictzwaarte.

Wat betreft de data governance-vraag bleek uit de focusgroepen dat er juridische en technische mogelijkheden zijn voor structurele levering van Ritax-data op microniveau aan het WODC. Vooralsnog ontbreekt een blijvende, inhoudelijke noodzaak die een structurele uitwisseling van dergelijke gegevens op microniveau vanuit de AVG rechtvaardigt. Structurele levering vraagt bovendien om aanzienlijke personele en materiële capaciteit, beveiligingsmaatregelen en heldere afspraken tussen de betrokken partijen. Daar is op dit moment nog geen voldoende noodzaak toe.

Deze haalbaarheidsstudie heeft enkele beperkingen en biedt tevens aanknopingspunten voor vervolg. De empirische analyses bestrijken slechts twee recente jaren en bieden daarmee geen zicht op ontwikkelingen over de wat langere termijn. Daarnaast blijft de selectie van jongeren waarvan Ritax-data zijn geregistreerd onduidelijk, waardoor de representativiteit van de gegevens ten aanzien van jeugdige verdachten onzeker is. Het is in deze studie niet duidelijk geworden waarom bij een aanzienlijk deel van de variabelen 'onbekend' wordt gescoord, waardoor de oorzaken van deze ontbrekende waarden onduidelijk blijven. Wel wijzen kwalitatieve bevindingen in de richting van variatie in afname en moeilijk te verkrijgen informatie voor bepaalde items. Desondanks biedt deze studie belangrijke aanknopingspunten. Voor het eerst is inzichtelijk gemaakt hoe consistent en volledig de Ritax-data op landelijk niveau zijn en waar verschillen zitten tussen domeinen, regio's en subgroepen. Als de betrouwbaarheid en daarmee de bruikbaarheid van de Ritax-data in de toekomst verbetert, kunnen deze gegevens bijdragen aan

evaluatieonderzoek, subgroep vergelijkingen en duiding van trends in jeugdcriminaliteit.

Alles overziend blijkt uit deze haalbaarheidsstudie dat het op dit moment niet voldoende verantwoord is om Ritax-data structureel te gebruiken voor periodiek onderzoek naar veranderingen over de tijd in achtergrondkenmerken van de totale populatie jeugdige verdachten en regionale en subgroep variatie daarin. Ofwel, tweejaarlijks monitoren zoals in de MJC ligt momenteel niet voor de hand. Vooralsnog vertonen de gegevens te veel variatie in registratie en invulling om betrouwbare vergelijkingen te maken. Als de datakwaliteit wordt versterkt, zou Ritax-data op termijn wél een betrouwbare bron kunnen worden voor dergelijke onderzoeksdoeleinden. Om de kwaliteit en bruikbaarheid van gegevens uit de Ritax te vergroten, is met name het terugdringen van 'onbekend'-scores in de domeinen met veel ontbrekende waarden van belang. Gedacht kan worden aan meer uniformiteit in afname- en registratie werkwijzen in de praktijk. Uiteraard moet hierbij ook rekening worden gehouden met de discretionaire ruimte van raadsonderzoekers zodat in de individuele onderzoeken ook maatwerk kan worden geleverd. Daarnaast is systematische documentatie van inhoudelijke wijzigingen van het instrument essentieel om mogelijk toekomstige methodebreuken te kunnen duiden. Training van raadsonderzoekers en verscherpte landelijke afstemming in registratierichtlijnen kan bijdragen aan consistentere en betrouwbaardere gegevens zodat deze ook bruikbaar worden om veranderingen in de populatie jeugdige verdachten over de tijd te onderzoeken.

Incidenteel onderzoek naar veranderingen in achtergrondkenmerken van jeugdige verdachten of veroordeelden over de tijd waarbij Ritax-data over meerdere jaren worden gebruikt lijkt wel uitvoerbaar, mits aandacht wordt besteed aan de in dit onderzoek geschetste beperkingen. Er dient dan wel zorgvuldig te worden afgewogen hoe in de onderzoeksopzet om te gaan met de mate van ontbrekende waarden en de gevolgen daarvan voor de conclusies van dergelijk onderzoek. Om de selectiviteit van de informatie te beperken en uitspraken over de populatie van jeugdige verdachten te kunnen doen, is op microniveau een koppeling nodig aan een externe bron waarin gegevens over de volledige populatie van jeugdige verdachten of justitiabelen beschikbaar zijn (denk aan de politieregistratie bij het CBS of de OBJD bij het WODC).

1 Inleiding

Met de Monitor Jeugdcriminaliteit (MJC) wordt periodiek (tweejaarlijks) onderzoek gedaan naar ontwikkelingen en veranderingen in jeugdcriminaliteit onder de algemene populatie van jeugdigen (12-23 jaar), jeugdige verdachten en jeugdige veroordeelden (WODC, n.d.). Om een breed en genuanceerd beeld te krijgen van dergelijke trends worden in de MJC diverse databronnen gecombineerd. Bijvoorbeeld politie data, justitiële data en zelfrapportage data (Van der Laan et al., 2024). Met de zelfrapportagedata kunnen bijvoorbeeld veranderingen over de tijd in risico- en beschermende factoren van jeugdigen worden beschreven. Dit gebeurt op basis van een grote, representatieve steekproef (Tollenaar et al., 2024). Desalniettemin betreft de zelfrapportagedata een steekproef uit de algemene populatie van jeugdigen. Jeugdigen die zijn verdacht door de politie zijn daarin wel vertegenwoordigd, maar het betreft niet alle jeugdige verdachten. Aanvullende databronnen van politie of justitie die ook gegevens over risico en beschermende factoren van jeugdige verdachten of justitiabelen registreren, zouden een waardevolle aanvulling kunnen vormen. Dit kan helpen bij een betere duiding van trends in geregistreerde jeugdcriminaliteit over de tijd. Bovendien klinken er geregeld geluiden uit de politie- en justitiepraktijk over een toenemende complexiteit van strafzaken en de onderliggende problematiek bij jeugdige justitiabelen. Als er inderdaad sprake is van een toegenomen complexiteit van onderliggende problematiek, dan zouden justitiële jeugdinterventies en zorg mogelijk moeten worden uitgebreid of aangepast.

Echter, veranderingen in de complexiteit van een specifieke doelgroep kunnen alleen worden vastgesteld op basis van een beschikbaar instrument. Een dergelijk instrument moet risico- en beschermende factoren van jeugdige justitiabelen consequent én consistent over de tijd meten en registreren. In het Nederlandse jeugdstrafrechtstelsel bestaat sinds 2009 het Landelijk Instrumentarium Jeugdstrafrecht (LIJ). Sinds 2013 wordt dit instrument structureel toegepast bij jeugdigen die met politie of justitie in aanraking komen (Van der Put et al., 2011; Van der Put & Kapteijns, 2020). Het LIJ is een samenhangend pakket van screeningsinstrumenten dat door diverse actoren in de jeugdstrafrechtstelsel wordt gebruikt, zoals de politie, Bureau Halt, de Raad voor de Kinderbescherming (RvdK) en jeugdreclassering (Buysse & Petersen, 2023; Ministerie van Justitie en Veiligheid (Ministerie van JenV), 2012; Van der Put et al., 2011). Het LIJ omvat onder andere: het Preselect Recidive, dat het algemeen recidiverisico bepaalt, het Halt signaleringsinstrument, en de Ritax, welke een screeningsinstrument betreft.

Het LIJ biedt mogelijk de gewenste informatie die nodig is om meer achtergrond en duiding te geven aan ontwikkelingen in geregistreerde jeugdcriminaliteit over de tijd. Geregistreerde LIJ-gegevens bevatten namelijk specifieke informatie over jeugdige verdachten en justitiabelen en vullen daarmee gegevens over jongeren uit de algemene populatie aan. Om verdiepend onderzoek te verrichten naar bijvoorbeeld veranderingen in kenmerken van de populatie van jeugdige verdachten of justitiabelen is het van belang dat gegevens uit het LIJ op microniveau beschikbaar zijn. Ook biedt dit mogelijkheden voor onderzoek naar effecten van interventies. Naast de mogelijkheden om veranderingen in de doelgroep over de tijd te beschrijven, zou dit dan ook kansen bieden om met observationele data quasi-experimenteel onderzoek naar effecten van justitiële jeugdinterventies uit te voeren. De data bieden namelijk mogelijkheden om controlegroepen met verschillende risicoprofielen samen te stellen.

Als LIJ-gegevens voldoende valide en betrouwbaar worden gemeten en geregistreerd over de tijd, kunnen deze worden gebruikt in de MJC om meer inzicht te verschaffen in ontwikkelingen in de geregistreeerde jeugdcriminaliteit, met name wat betreft jeugdige verdachten. Op dit moment is het echter onbekend of de gegevens over de tijd dusdanig betrouwbaar zijn geregistreerd zodat ze bruikbaar zijn als aanvullende bron voor de Monitor Jeugdcriminaliteit (MJC). Er zijn namelijk wel aanwijzingen dat het LIJ sinds de landelijke invoering ervan in 2013 veranderd is, met mogelijke gevolgen voor de betrouwbaarheid van de gegevens voor vergelijkingen over de tijd. Zo heeft het instrument recentelijk een uitgebreide aanpassing ondergaan, waarbij zowel items binnen domeinen als antwoordcategorieën binnen bepaalde items zijn aangepast of verwijderd (Van der Put & Kapteins, 2020). Daarnaast worden in recent onderzoek organisatorische problemen benoemd, zoals personeelsverloop en regionale verschillen daarin en problemen met het registratiesysteem zoals moeilijkheden in de werkbaarheid van het rapportageformat en een gebrekkige koppeling tussen verschillende soorten registratiesystemen (Buysse & Petersen, 2023). Deze en andere veranderingen, evenals problemen in de afname en registratie van de gegevens, roepen vragen op over de betrouwbaarheid van gegevens uit het LIJ om veranderingen in de doelgroep over de tijd te meten.

Het is op dit moment onduidelijk of gegevens uit het LIJ geschikt zijn om vergelijkende analyses over de tijd uit te voeren. Vanuit zowel maatschappelijk als wetenschappelijk oogpunt is het belangrijk om betrouwbare gegevens te hebben voor het beschrijven en duiden van eventuele veranderingen in de doelgroep van jeugdige verdachten of justitiabelen. Daarom is verkennend onderzoek naar de betrouwbaarheid en consistentie van geregistreeerde gegevens uit het LIJ over de tijd noodzakelijk. Een haalbaarheidsstudie moet daarom eerst plaatsvinden voordat deze data structureel gebruikt kunnen worden om ontwikkelingen onder (deel)populaties van jeugdige verdachten of justitiabelen over meerdere jaren in kaart te brengen. In dit rapport worden de resultaten van deze haalbaarheidsstudie beschreven. Het onderzoek is beperkt tot gegevens die via de RvdK worden verzameld en betreft daarmee specifiek de Ritax A en B die worden gebruikt door de RvdK. Het gaat niet over andere delen van het LIJ (zoals het Preselect Recidive), van andere ketenpartners, of gerelateerde screeningsinstrumenten zoals de Halt SI. Zie ook paragraaf 2.1.

Dit onderzoek gaat over de bruikbaarheid van de Ritax A en B van de RvdK, *voor vergelijkingen over de tijd op (sub)groep- of populatieniveau*. Het gaat niet over het gebruik ervan voor individuele (strafrechtelijke) trajecten of de mate waarin de gegevens recidive voorspellen. Daar is reeds eerder uitgebreid onderzoek naar gedaan. Waar nodig wordt dergelijk eerder onderzoek wel aangehaald of kort besproken.

1.1 Achtergrond

1.1.1 Beleidscontext

Voor het beleid van het ministerie van JenV is het relevant om de ontwikkelingen in de jeugdcriminaliteit in Nederland periodiek te kunnen volgen, daarover aan de TK te rapporteren (zie bijv. *Kamerstukken II*, 2022/23, 28741, nr. 107) en de achtergronden van die ontwikkelingen te kennen. Vanuit het ministerie van JenV is de vraag gekomen of het LIJ als aanvullende bron binnen de MJC kan worden ingezet om meer duiding en achtergrond te geven in de ontwikkelingen en veranderingen in jeugdcriminaliteit in de

loop der tijd. Daarnaast zijn er momenteel verschillende geluiden uit de praktijk over een toegenomen complexiteit van onderliggende problematiek onder jeugdige justitiabelen. Inzicht in ontwikkelingen in risicoprofielen van jeugdige justitiabelen over de tijd kunnen bijdragen aan de kennis over de achtergrond van ontwikkelingen in de (populatie) jeugdcriminaliteit. Dit is relevant voor beleidsmakers en professionals in de jeugdstrafrechtketen, die het bijvoorbeeld kunnen gebruiken om beter geïnformeerde beleidsbeslissingen te nemen in de jeugdstrafrechtketen en waar nodig, aan te passen om beter tegemoet te komen aan risicoprofielen van jeugdige justitiabelen. Gegevens van het LIJ zouden inzicht kunnen geven in dit soort ontwikkelingen en achtergronden. Echter, alvorens het gebruik van gegevens uit het LIJ voor dit doeleinde mogelijk is, is het belangrijk om inzicht te krijgen in de betrouwbaarheid van de registratie van deze gegevens over de tijd. Als blijkt dat de data over de tijd op populatieniveau (en voor subgroepen daarbinnen) voldoende betrouwbaar is geregistreerd, kan dit mogelijkheden bieden om in de toekomst meer inzicht te krijgen in trends in jeugdcriminaliteit, alsmede de effectiviteit van interventies en verschuivingen in de profielen van jeugdige justitiabelen te onderzoeken.

1.1.2 Doel- en vraagstelling

De doelstelling van deze haalbaarheidsstudie is tweeledig. Allereerst is het overkoepelende doel te beoordelen in hoeverre gegevens uit het LIJ bruikbaar zijn om veranderingen over de tijd in risico- en beschermende factoren van (deel)populaties jeugdige verdachten te meten. Ten tweede is het zo dat gegevens vanuit het LIJ in de afgelopen jaren herhaaldelijk zijn gebruikt in incidentele studies (Buysse & Petersen, 2023; Buysse et al., 2022; Egberink & Meijer, 2012; Mensink et al., 2021; Nauta & Loef, 2011; Timmermans & Witvliet, 2011; Van der Ark et al., 2018; Van der Put & Kapteijns, 2020). Telkens opnieuw leidde dit tot langdurige voorbereidende trajecten, onder meer vanuit privacy- en data governance-perspectief. Gezien het feit dat er in de afgelopen jaren meerdere studies zijn uitgevoerd, ligt het voor de hand om na te denken over een structurele levering van deze gegevens op microniveau aan het WODC voor verdiepend wetenschappelijk onderzoek. Echter, structurele beschikbaarheid over (een afslag van) het LIJ op microniveau is alleen mogelijk als aan diverse randvoorwaarden is voldaan, zoals voldoende noodzaak om structureel LIJ-data beschikbaar te hebben binnen het WODC (grondslag). Daarnaast moet ook worden nagedacht over de privacyvoorwaarden, evenals organisatorische en logistieke mogelijkheden voor onder meer dataverwerking en -opslag. Het tweede doel van dit onderzoek is daarom om te verkennen aan welke voorwaarden met betrekking tot onder meer privacy, data governance, en logistiek moet worden voldaan om op structurele basis over LIJ-data op microniveau te beschikken voor verdiepend onderzoek.

Als gevolg van deze tweeledige doelstelling wordt deze haalbaarheidsstudie opgesplitst in twee onderdelen, namelijk een (I) betrouwbaarheidsvraag en een (II) vraag omtrent data governance, waarbij de tweede afhankelijk is van het antwoord op de eerste vraag.

De centrale vraag betreffende de betrouwbaarheid van de data voor metingen over de tijd is:

In hoeverre zijn gegevens uit het LIJ geschikt om ontwikkelingen te meten met betrekking tot geaggregeerde trends van risico- en beschermende factoren van (deel)populaties minderjarige verdachten?

Om de bruikbaarheid van gegevens uit het LIJ over de tijd te kunnen beoordelen, is het allereerst belangrijk om te kijken naar veranderingen in de registratie- of meetmethoden van het LIJ sinds de invoering ervan in 2013. Daarnaast moet, om ontwikkelingen vast te kunnen stellen, de vulling van de individuele gegevens volledig en consistent zijn door de tijd heen. Daarom is het tevens belangrijk om te bestuderen in hoeverre de gegevens over de tijd voldoende ingevuld zijn, of juist problematische missende gegevens bevatten. In de methode sectie definiëren we dit nader. Om te beoordelen of gegevens uit het LIJ gebruikt kunnen worden om veranderingen in de populatie van jeugdige justitiabelen over de tijd te kunnen beschrijven en duiden, zou dan, voor de consistent geregistreerde gegevens, ook exploratief gekeken kunnen worden naar veranderingen over de tijd wat betreft demografische of delict kenmerken. Vervolgens zou bekeken moeten worden welke maatregelen mogelijk genomen moeten worden om gebreken op te vangen.

Daarom staan in de (I) betrouwbaarheidsvraag de volgende onderzoeksvragen centraal:

- 1 a In hoeverre worden gegevens uit het LIJ door de tijd heen op dezelfde manier geregistreerd?
b Wanneer treden er veranderingen op in de manier van meten, waar, wanneer en waarom treden deze veranderingen op?
- 2 a In welke mate zijn de gegevens uit het LIJ voldoende (in)gevuld over de tijd?
b In hoeverre zijn er verschillen tussen de domeinen en items in de mate waarin gegevens ontbreken?
- 3 Vanaf welke periode zijn gegevens uit het LIJ geschikt om te gebruiken?
- 4 Welke veranderingen zijn er in de populatie van jeugdige justitiabelen met een LIJ-afname door de tijd heen wat betreft demografische gegevens (zoals sekse, leeftijd, regio) en delict kenmerken (type delict, delict zwaarte)?
- 5 Welke aanpassingen zouden noodzakelijk en tevens haalbaar zijn om de kwaliteit van geregistreerde gegevens uit het LIJ, indien nodig, te verbeteren?

Zoals eerder opgemerkt gaat het bij het empirische deel van dit onderzoek specifiek om de Ritax-data van de RvdK, welke onderdeel zijn van het LIJ.

Als de eerste vraag naar consistente registratie van de gegevens over de tijd, de betrouwbaarheidsvraag, positief wordt beantwoord, rijst de vervolgvraag of de microgegevens op structurele basis door de bronbeheerder of via het Directoraat Generaal Straffen en Beschermen (DGSenB) Ketenregie aan het WODC kunnen worden geleverd voor wetenschappelijk onderzoek. De microdata kunnen dan niet alleen voor de MJC worden ingezet, maar ook voor criminele carrièreonderzoek en evaluatieonderzoek van straffen en interventies. Verschillende data governance-aspecten zijn hierin van belang om nader te onderzoeken. Het gaat bijvoorbeeld om privacy vragen omtrent het vaststellen welke stappen ondernomen moeten worden om AVG-compliant een dergelijke uitwisseling van microdata te realiseren en of deze uitvoerbaar zijn. Daarnaast zijn er ook organisatorische aspecten die moeten worden verkend, zoals de relatie tussen en bereidwilligheid van de betrokken partijen (d.w.z., WODC, RvdK en DGSenB) om een dergelijke uitwisseling te faciliteren. In het geval van een dergelijke uitwisseling in de toekomst, moet gekeken worden in hoeverre het

haalbaar is, vanuit privacy, organisatorisch, logistiek, budgettair en data governance-perspectief, om op structurele basis LIJ-data te huisvesten binnen het WODC. Hier dient dan ook aandacht te worden besteed aan logistieke aspecten, met name wat betreft de (technische) mogelijkheden en behoeften om dergelijke data op structurele wijze periodiek uit te wisselen. Tot slot is er een budgettaire kwestie omtrent de financiële middelen die nodig zijn om de voorgestelde werkzaamheden te financieren en daaropvolgend een zorgvuldige kosten-baten afweging om te bepalen of het mogelijk is om dergelijke data binnen het WODC te huisvesten. Tijdens de dataverzameling in dit onderzoek bleek dat de data governance-onderzoeksvragen (zie vragen 6 t/m 9)¹ pas goed beantwoord kunnen worden zodra de eerste betrouwbaarheidsvraag van dit onderzoek volledig is afgerond. Daarom worden de vragen 6 t/m 9 in dit onderzoek niet in deze haalbaarheidsstudie beantwoord. Wel is onderzocht wat de mogelijkheden en beperkingen zijn omtrent het structureel leveren van LIJ-data aan het WODC, vanuit de databasebeheerder (Justid), de dataeigenaar (RvdK) en DGSenB Ketenregie.

1.2 Leeswijzer

In hoofdstuk 2 worden de methoden van dit onderzoek omschreven. In hoofdstuk 3 worden de resultaten van de betrouwbaarheidsvraag besproken. Allereerst komen in paragraaf 3.1 de resultaten van de literatuurstudie aan bod, met daaropvolgend de resultaten van het kwalitatieve gedeelte in paragraaf 3.2, en de resultaten van het kwantitatieve gedeelte in paragraaf 3.3. In paragraaf 3.4 worden de resultaten van de interviews en focusgroepen omtrent de data governance-vraag besproken. Afsluitend bevat hoofdstuk 5 een slotbeschouwing en samenvatting van het onderzoek.

¹ Oorspronkelijk geformuleerde data governance-onderzoeksvragen die in deze haalbaarheidsstudie niet worden beantwoord:

- 6a Is er een grondslag voor het structureel beheren van LIJ-data binnen het WODC?
- 6b Welke maatregelen zijn er nodig om de privacy-risico's van structureel beheer binnen het WODC te mitigeren?
- 6c Is het beheer alleen voor intern WODC-gebruik of betreft het structureel beheer ook verstrekking van gegevens aan derden?
- 7a Welke taken en verantwoordelijkheden omvat het structureel beheer van LIJ-data binnen het WODC?
- 7b Waar kunnen de taken en verantwoordelijkheden voor het structureel beheer van LIJ-data het best worden belegd?
- 8a Hoe kan de logistieke inrichting voor transport, opslag, beheer en gebruik van LIJ-data optimaal het beste worden vormgegeven?
- 8b Welke infrastructuur is er nodig en hoe kan deze ingericht worden?
- 9a Wat zijn de financiële implicaties van het huisvesten van LIJ-data binnen het WODC?
- 9b Wat zijn de mogelijkheden voor bekostiging?

2 Methoden

In dit hoofdstuk worden de methoden beschreven die gehanteerd zijn in deze haalbaarheidsstudie. Hierin wordt gebruikgemaakt van een mixed methods design, omdat zowel kwalitatieve als kwantitatieve methoden nodig zijn om de relevante onderzoeksvragen omvattend te beantwoorden. Hieronder worden eerst de afbakeningen binnen de methoden van deze haalbaarheidsstudie beschreven. Vervolgens wordt het gebruik van elke methode toegespitst op de specifieke onderzoeksvragen met betrekking tot allereerst de betrouwbaarheidsvraag en daarna de data governance-vraag.

2.1 Afbakeningen

De validiteit van gegevens uit het LIJ, oftewel de mate waarin het instrument daadwerkelijk meet wat het beoogd te meten, is al vaker onderzocht in verschillende (recente) studies (zie o.a. Egberink & Meijer, 2012; Timmermans & Witvliet, 2011; Van der Ark et al., 2018; Van der Put & Kapteijns, 2020). Daarom richt deze haalbaarheidsstudie zich specifiek op het onderzoeken van de *betrouwbaarheid* van de gegevens uit het LIJ voor vergelijkingen over de tijd op populatieniveau of bij subgroepen. Namelijk, de mate waarin afnames consistent en herhaalbaar zijn uitgevoerd over de tijd, waarbij we ook onderzoeken of er landelijk op vergelijkbare manier wordt geregistreerd.

Met 'betrouwbaarheid' wordt hier bedoeld op de interne consistentie in mate en wijze van registratie van gegevens gemeten op jaarbasis, waarbij ook gekeken wordt naar eventuele geografische variaties. Daarbij wordt eerst, onder andere, de consistentie onderzocht in de wijze waarop onderzoeken in de praktijk worden afgenomen en geregistreerd in databases. Dit is een kwalitatief onderzoek op basis van literatuurstudie en gesprekken met professionals.

Vervolgens wordt kwantitatief de mate van invulling van geregistreerde gegevens in bestaande datasets over meerdere jaren onderzocht. Er wordt gekeken naar de mate van problematische ontbrekende waarden (zoals lege velden of 'onbekend' scores) en opvallende variaties daarin op domein- en itemniveau. Vanaf welk percentage ontbrekende waarden als 'problematisch' wordt beschouwd, kan op twee manieren worden bekeken: vanuit de praktijk of vanuit wetenschappelijk oogpunt. In de praktijk wordt doorgaans een drempel van 20% gehanteerd, waarbij variabelen (de items) met meer dan 20% ontbrekende waarden over de hele groep als problematisch worden beschouwd (P. 845, persoonlijke communicatie, 26 maart 2024). Vanuit wetenschappelijk oogpunt variëren acceptabele drempelwaardes tussen de 5-20%, waarbij variabelen met minder dan 5% ontbrekende waarden geen extra aandacht vereisen, variabelen met 5-20% ontbrekende waarden voorzichtig moeten worden behandeld (en moet worden overwogen hoe om te gaan met de ontbrekende waarden), en voor variabelen met meer dan 20% ontbrekende waarden exclusie in overweging moet worden genomen (Enders, 2022). Bij de kwantitatieve beoordeling van de betrouwbaarheid van de gegevens voor vergelijkingen over de tijd wordt in deze haalbaarheidsstudie rekening gehouden met zowel de praktische als de wetenschappelijke drempelwaarden. Verder wordt ook geëvalueerd in hoeverre het LIJ in de loop der tijd is veranderd en wat de consequenties hiervan zijn voor de

vergelijkbaarheid en bruikbaarheid van de gegevens over de tijd. Voor een overzicht van de afbakeningen met betrekking tot betrouwbaarheid, zie tabel 2.3.

Zoals in de inleiding benoemd, richt het empirische onderzoek zich uitsluitend op een selectie van gegevens uit het LIJ, namelijk de Ritax A (voorheen instrument 2a) en Ritax B (voorheen instrument 2b) instrumenten. Deze gegevens zijn eigendom van de RvdK en kunnen worden verkregen via Justid, gecoördineerd door DGSenB. De keuze om enkel een deel van gegevens uit het LIJ te gebruiken heeft te maken met het feit dat verschillende organisaties eigenaar zijn van afzonderlijke delen. Het gebruik van LIJ-data uit alle delen van de jeugdstrafrechtketen zou complexe en langdurige toestemmingsprocedures vereisen bij meerdere broneigenaren. Het gaat daarbij om onder andere de politie, bureau Halt, de gecertificeerde instellingen, de RvdK en justitiële jeugdinrichtingen. Gezien de verkennende aard van deze haalbaarheidsstudie zou een selectie toereikend zijn.

2.2 Methode

2.2.1 Betrouwbaarheid

Er zijn verschillende onderzoeksmethoden gebruikt om antwoord te krijgen op de mate waarin de gegevens over de tijd betrouwbaar zijn geregistreerd. Naast een literatuurstudie verrichten we ook kwalitatief empirisch onderzoek door interviews en focusgroepen met experts, en kwantitatief empirisch onderzoek door gegevens van de Ritax A en Ritax B te analyseren.

Literatuurstudie

Ten behoeve van de betrouwbaarheidsvraag is allereerst een literatuurstudie en analyse uitgevoerd om onderzoeksvragen 1a, 1b, 2a, 2b, en 3 te beantwoorden. De bevindingen uit relevante, zowel gepubliceerde als ongepubliceerde literatuur zijn geanalyseerd om te bepalen wat er al bekend is over de (validiteit en) betrouwbaarheid van gegevens uit het LIJ. Daarnaast is gekeken naar de constructie van het instrument en eventuele aanpassingen door de tijd heen. Ook is gekeken naar wat er al bekend is met betrekking tot de mate van invulling van gegevens, en mogelijke verschillen tussen domeinen en items, om inzicht te verkrijgen in vanaf welke periode de gegevens eventueel geschikt zou zijn voor gebruik binnen de MJC. De initiële zoektocht richtte zich op onderzoeken naar de LIJ-instrumenten binnen de Nederlandse jeugdstrafrechtketen, uitgevoerd door of in opdracht van het WODC. Daarom begon de literatuurzoektocht in de WODC Repository.

Om te voorkomen dat overige onderzoeken, welke niet in opdracht van het WODC zijn uitgevoerd, gemist werden, is in de tweede plaats ook in bredere academische databases gezocht. Dit betrof Web of Science en Google Scholar. Hierbij is gebruikgemaakt van Nederlandstalige en Engelstalige trefwoorden die gerelateerd zijn aan het doel en de focus van dit onderzoek, in combinatie met Booleaanse operatoren.² De Nederlandstalige termen van deze zoekopdracht waren: 'LIJ OF Landelijk Instrumentarium Jeugdstrafrechtketen OF risico* OF instrument* EN validiteit OF betrouwbaarheid EN/OF jeugd* OF jong* OF kind*'. De Engelstalige termen van deze zoekopdracht zijn: 'LIJ OR Landelijk Instrumentarium

² Booleaanse operatoren zijn logische operatoren -zoals bijvoorbeeld AND, OR, en NOT- die gebruikt worden om zoekopdrachten preciezer te maken en irrelevante informatie te filteren in data-analyse en informatiesystemen (Machi & McEvoy, 2016).

Jeugdstrafrechtketen OR National Risk Assessment Tool OR Risk* OR instrument* AND Netherlands OR NLs OR Dutch OR Holland AND validity OR reliability OR effect* AND/OR youth* OR juvenile* OR young* OR child*’.

Ten derde is de zoekactie verder uitgebreid met de sneeuwbalmethode: hierbij zijn referenties binnen de gevonden publicaties nagegaan op relevante titel. Ook zijn auteurs van belangrijke studies gecontacteerd om aanvullende werken te vinden die niet via trefwoordzoekacties naar voren zijn gekomen. Het inclusiecriteria (en tevens exclusie-) was dat studies expliciet betrekking moesten hebben op analyses van gegevens uit het LIJ. Om ervoor te zorgen dat geen waardevolle informatie over het hoofd werd gezien, is daarnaast actief contact opgenomen worden met relevante medewerkers van betrokken organisaties. Bijvoorbeeld, de RvdK, het WODC en Ketenregie, met het verzoek om inzage in relevante grijze literatuur zoals ongepubliceerde beleidsdocumenten die inzicht geven in de wijze van afname en registratie in de praktijk.

Interviews en focusgroepen

Ter verdere beantwoording van onderzoeksvragen 1 tot en met 3, en gedeeltelijk onderzoeksvragen 4 en 5, zijn expertinterviews en focusgroepen gehouden. Beide methoden zijn gekozen om verschillende perspectieven te verzamelen. De interviews boden diepgaande, individuele inzichten. De focusgroepen vormden een interactieve omgeving boden voor het verzamelen van veelvuldige en meerzijdige perspectieven van experts uit de praktijk. De interviews en focusgroepen duurden ongeveer één uur, en waren semigestructureerd van aard. Dit betekende dat er vooraf gesprekstema's waren vastgesteld, en tevens ruimte was voor spontaniteit en het verkennen van nieuwe onderwerpen binnen deze thema's, afhankelijk van de reacties van de deelnemers. Interview deelnemers zijn gerekruteerd middels een doelgerichte steekproef (*purposive sampling*), waarbij is gestreefd maximale variatie onder respondenten. Hierbij lag de focus op 'LIJ-experts' met ruime ervaring vanuit beleid, onderzoek, of praktijkperspectief met betrekking tot de registratie en afname van instrumenten uit het LIJ over de tijd. Dit omvatte met name wetenschappelijke experts, zoals onderzoekers die validatiestudies hebben uitgevoerd. Ook beleidsexperts waren geïnccludeerd, waaronder senior beleidsmedewerkers met inzicht in de praktische toepassing van het LIJ en eventuele veranderingen daarin over de tijd. Volgens de standaarden voor kwalitatief onderzoek zijn experts geworven tot het punt van 'saturatie'. Dit houdt in dat er geen nieuwe inzichten of thema's meer naar voren kwamen uit de interviews. Daarnaast zijn er in totaal twee focusgroepen georganiseerd met praktijkexperts die ervaring hebben met het afnemen en registreren van onderzoeken in hun functie als raadsonderzoeker. Ook hier is gestreefd naar maximale variatie. De deelnemers verschilden zowel in de duur van hun werkervaring als hun werklocaties binnen Nederland, om een goede spreiding in zowel werkervaring en geografie te waarborgen.

In totaal hebben twintig (praktijk)experts hun medewerking verleend, waarvan tien experts deelnamen aan de expertinterviews en tien experts aan de focusgroepen. Alle experts waren medewerkers van het ministerie van JenV, van wie het merendeel werkzaam was bij de RvdK. Er was geen overlap tussen de deelnemers waardoor beide onderdelen unieke perspectieven boden. In tabel 2.1 staat een overzicht van de functiesoorten en het aantal jaren werkervaring van de interviewdeelnemers (hierna: respondenten). In tabel 2.2 staat een overzicht van de werkregio's en aantal jaren werkervaring van de focusgroepdeelnemers (hierna: respondenten), die waren verdeeld over twee afzonderlijke focusgroepen.

Tabel 2.1 Kenmerken interviewdeelnemers

Kenmerk	Aantal (N=10)
Functiesoort	
Onderzoeker	3
(LIJ)Data Officer/Specialist	5
Adviseur	3
Projectbegeleider	1*
Jaren werkervaring	
0 tot 5	3
5 tot 10	0
10 tot 15	2
15 of meer	5

* N>10 vanwege duale functierollen bij sommige participanten.

Tabel 2.2 Kenmerken focusgroepdeelnemers betrouwbaarheidsvraag

Kenmerk	Aantal (N=10)
Werkregio	
Noord-Holland	1
Gelderland	1
Midden-Nederland	5
Noord-Nederland	1
Oost-Brabant	1
West-Brabant	1
Jaren werkervaring	
0 tot 5	3
5 tot 10	3
10 tot 15	0
15 of meer	4

Kwantitatief onderzoek

Om onderzoeksvragen 2 tot en met 5 van de betrouwbaarheidsvraag verder te beantwoorden zijn ook kwantitatieve methoden ingezet. Dit omvatte het opvragen van geanonimiseerde datasets op itemniveau over meerdere jaren bij Justid via DGSenB en bij de RvdK. Deze datasets boden waardevolle inzichten in de mate van invulling van geregistreerde data over de tijd. Hierbij is vooral gekeken naar de volledigheid van (invulling van) gegevens over de tijd. Ook is gekeken naar mogelijke patronen in ontbrekende gegevens voor specifieke groepen, jaren of regio's, en binnen de verschillende items en domeinen. Zoals eerder vermeld in paragraaf 2.1, betrof dit uitsluitend gegevens van de Ritax A en Ritax B, waarvan de RvdK eigenaar is.

In eerste instantie is een dataset op itemniveau opgevraagd bij DGSenB Ketenregie, welke structureel de Ritax-data van de RvdK beschikbaar hebben in verband met

monitoring van de Ritax. Deze bevatte gegevens van 2020 tot en met oktober 2023. Het verkrijgen van oudere data op itemniveau (van vóór 2020) was op dat moment niet mogelijk. Dit kwam door technische beperkingen in de historische opslag, waardoor het verzamelen van oudere gegevens buiten het bereik van deze haalbaarheidsstudie viel. Gezamenlijk kunnen de datasets die wél werden verkregen een voorzichtige indicatie geven van trends en veranderingen in de ingevulde data over de tijd. Vervolgens is beoordeeld of reeds beschikbare data voldoende bruikbaar zijn voor gerichte, toekomstige analyses met een meerjarig perspectief, of wat er eventueel nodig is om de gegevens in de toekomst geschikt te maken voor dit doeleinde.

Daarvoor was het essentieel om condities te identificeren die van invloed kunnen zijn op de betrouwbaarheid van geregistreerde data voor vergelijkingen over de tijd. Het expliciet maken van deze condities was daarbij behulpzaam. Bij het beoordelen van de kwaliteit van de data over de tijd speelden drie groepen condities een belangrijke rol, namelijk: inhoudelijke wijzigingen van het instrument, ervaringen van professionals in de uitvoering en data-indicatoren voor consistentie en volledigheid in de invulling van geregistreerde gegevens. Deze condities zijn weergegeven in tabel 2.3.

Tabel 2.3 **Conditie voor de bruikbaarheid van LIJ-data voor vergelijkingen over de tijd**

Conditiesoort	Criteria
Inhoudelijke wijzigingen of knelpunten	Geen aanwijzingen uit de literatuur voor significante wijzigingen of knelpunten in: a) de inhoud van het LIJ b) gebruikte ICT-systemen c) wijze van afname
Ervaringen van professionals	Geen meldingen uit interviews en focusgroepen van: a) significante veranderingen (in afname-instructies) over de tijd b) ongewone regionale variatie c) systematisch ontbrekende gegevens
Indicatoren in de data	Geen resultaten uit analyses van LIJ-datasets die wijzen op: a) significante wijzigingen in items b) problematische ontbrekende waarden in specifieke domeinen of items

2.2.2 *Data governance*

De data governance-vraag wordt relevant als gegevens uit het LIJ vanaf een bepaald moment voldoende betrouwbaar zijn voor structureel monitoren van ontwikkelingen over de tijd. Ter beantwoording van een overkoepelende, grondslag en noodzaak vraag wat betreft data governance werd eerst een focusgroep gehouden met uitsluitend WODC-medewerkers. In deze focusgroep is ingegaan op de vraag of er voldoende noodzaak is om structureel gegevens uit het LIJ op microniveau beschikbaar te hebben binnen het WODC. Daarnaast was het belangrijk om inzichten te verkrijgen in mogelijkheden en beperkingen omtrent het organiseren van een mogelijke, toekomstige structurele levering van data aan het WODC. Voor het structureel binnenhalen van microdata gelden namelijk strenge eisen; een eenmalige onderzoeksaanvraag biedt geen voldoende juridische grondslag voor structurele verwerking van microdata. Daarvoor zou ook sprake moeten zijn van een

aantoonbare, structurele noodzaak die het blijvend gebruik van microdata rechtvaardigt. Bijvoorbeeld, wanneer de gegevens niet alleen nodig zijn voor eenmalige vragen over de populatie, maar ook essentieel blijken voor het doorlopend monitoren van de populatie als relevante beleidsindicator (zoals bijvoorbeeld het geval is bij recidive), kan sprake zijn van een situatie met voldoende grondslag én noodzaak. Om dit onderwerp verder te onderzoeken is een tweede, externe focusgroep gehouden met belanghebbenden van buiten het WODC. Dit betrof data-professionals van DGSenB – die al een structurele dataleveringsovereenkomst hebben – en van Justid, als dataleveranciers van het ministerie van JenV. Beide focusgroepen duurden ongeveer één à anderhalf uur en hadden een semigestructureerd karakter. Voor beide focusgroepen is gewerkt met een doelgerichte steekproef. Deelnemers waren uitgenodigd op basis van hun ervaringen met datamanagement vraagstukken of (structurele) dataleveringen. Conform standaarden voor kwalitatief onderzoek bestonden de focusgroepen uit minimaal zes tot maximaal tien deelnemers.

In totaal namen twaalf medewerkers van het ministerie van JenV deel aan de twee focusgroepen. Zes namen deel aan de interne focusgroep en zes aan de externe focusgroep. Er was geen overlap tussen deze deelnemers. Medewerkers die aan de interne focusgroep deelnamen waren geen deelnemers van de externe focusgroep en vice versa. In tabel 2.4 staat een overzicht van de organisaties en de functiesoorten van de deelnemers.

Tabel 2.4 Kenmerken focusgroepdeelnemers data governance-vraag

Kenmerk	Aantal (N=12)
Functiesoort	
Privacy/Data Officer	2
Productspecialist	2
Data Architect	1
Adviseur	3
Projectbegeleider	1
Onderzoeker	3
Organisatie	
WODC	6
Justid	3
DGSenB	3

3 Resultaten

In dit hoofdstuk worden de resultaten van de huidige haalbaarheidsstudie beschreven. In paragraaf 3.1 worden de bevindingen van de literatuurstudie uiteengezet, ter beantwoording van onderzoeksvragen 1a, 1b, 2a, 2b en 3 betreffende de betrouwbaarheid van geregistreeerde LIJ-data over de tijd. In paragraaf 3.2 worden de resultaten van de interviews en focusgroepen welke ten behoeve van de betrouwbaarheidsvraag zijn afgenomen beschreven. Dit ter verdere beantwoording van vraag 1 t/m 3, 4 en 5. In paragraaf 3.3 worden de resultaten van de kwantitatieve analyses ten behoeve van de betrouwbaarheidsvraag en verdere beantwoording van onderzoeksvragen 2a, 2b en 3 beschreven. Paragraaf 3.4 omschrijft de resultaten van de focusgroepen ten behoeve van de data governance-vraag, waarin besproken is wat de mogelijkheden en beperkingen zijn omtrent een structurele levering van gegevens uit het LIJ aan het WODC.

3.1 Literatuurstudie

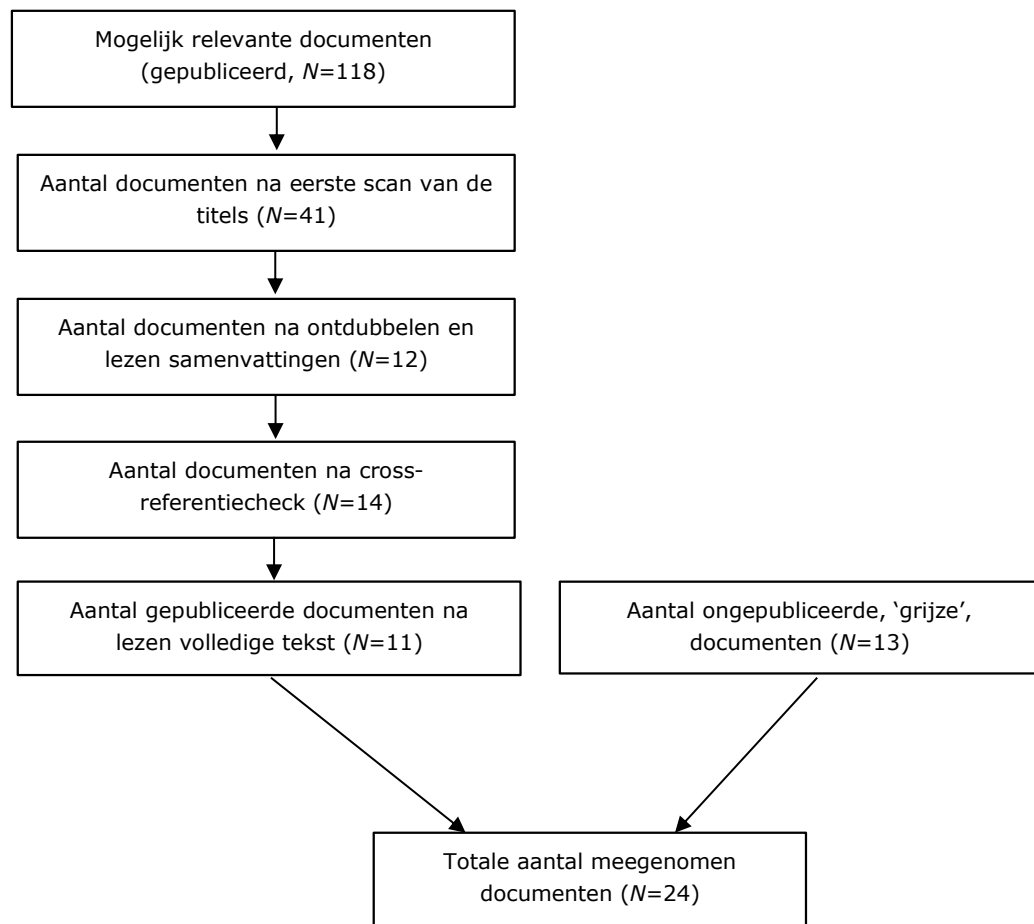
We beschrijven eerst de wijze waarop de literatuur is geselecteerd, geven vervolgens een korte beschrijving van het LIJ. Daarna gaan we specifiek in op bevindingen over validiteit, betrouwbaarheid en veranderingen in het LIJ over de tijd. In de tussenconclusie gaan we op basis van de literatuur in op de bruikbaarheid van het LIJ voor onderzoek naar ontwikkelingen in de jeugdige verdachten populatie over de tijd.

3.1.1 Zoekopdracht en documentselectie

In het kader van de literatuurstudie is een uitgebreide zoekopdracht uitgevoerd naar gepubliceerde documenten, waarbij gebruik is gemaakt van diverse trefwoorden in vooraf geselecteerde databases (zie paragraaf 2.2). De focus lag specifiek op literatuur die betrekking heeft op de betrouwbaarheid van het LIJ, zowel in theorie als in de praktijk van afname. Het inclusiecriterium (en tevens exclusiecriterium) was dat studies expliciet betrekking moesten hebben op de kwaliteit, validiteit of betrouwbaarheid van het LIJ. In eerste instantie resulteerde deze zoekopdracht in 118 mogelijk relevante documenten. Na een eerste scan van de titels werden 41 documenten geschikt bevonden voor inclusie, en na het lezen van de samenvattingen en ontdubbelen van deze documenten bleven 12 unieke publicaties over. Een aanvullende cross-referentiecheck leidde tot de identificatie van nog eens 2 documenten, waardoor het totale aantal op 14 kwam. De volledige teksten van deze 14 documenten zijn gelezen, waarna 3 studies alsnog werden uitgesloten wegens een gebrek aan relevantie, omdat zij betrekking hadden op een ander (risicotaxatie-)instrument dan het LIJ. Uiteindelijk zijn 11 gepubliceerde studies opgenomen in de huidige literatuurstudie. Tevens zijn verschillende medewerkers binnen het ministerie van JenV benaderd met het verzoek om ongepubliceerde, oftewel 'grijze', documenten aan te leveren. De medewerkers die benaderd werden moesten werkzaam zijn bij een ketenpartner die met het LIJ werkt en in hun functie op enige manier betrokken zijn (geweest) bij de ontwikkeling of uitvoering ervan. Deze uitvraag resulteerde in 13 aanvullende stukken, variërend van interne beleidsstukken (memo's, grafieken) tot handleidingen voor het gebruik en de afname van het LIJ in de praktijk. Daarnaast omvatte het een logboek met een overzicht van aanpassingen aan de dataset vanaf

2019 t/m heden, zoals geregistreerd door datamedewerkers van DGSenB. In totaal zijn dus 24 documenten meegenomen in dit literatuuronderzoek.

Figuur 3.1 Documenteselectieproces in de literatuurstudie



3.1.2 De samenstelling van het LIJ-instrument

Het LIJ is een risicotaxatie-instrument dat is ontworpen voor jeugdigen tussen de 12 tot 18 jaar die in aanraking komen met de jeugdstrafrechtketen, en moet worden toegepast in alle strafzaken (Ministerie van JenV, 2012). Sinds de invoering van het adolescentiestrafrecht in 2014 wordt het LIJ door de jeugdreclassering ook toegepast in strafzaken van adolescenten tussen de 18-23 jaar die middels het jeugdstrafrecht worden berecht (Buysse et al., 2022). Samenvattend biedt het LIJ een inventarisatie van risico- en beschermende factoren en zorgsignalen, die gezamenlijk leiden tot de samenstelling van een (recidive)risicoprofiel van jeugdige verdachten. Op basis van dit risicoprofiel kan worden aanbevolen welke strafrechtelijke benadering en eventuele zorg het meest passend zou zijn om de kans op recidive van de desbetreffende jongere te verkleinen (Ministerie van JenV, 2012; Ministerie van JenV, 2024).

Het LIJ bestaat uit verschillende instrumenten die worden afgenomen in alle fases van de jeugdstrafrechtketen. In eerste instantie wordt bij jeugdigen die in aanraking komen met de politie het 'Preselect Recidive'-instrument gebruikt, waarbij een

inschatting wordt gemaakt van het recidiverisico op basis van statische risicofactoren. Bij doorverwijzing naar Bureau Halt kan vervolgens het Halt-signaleringsinstrument (Halt-SI) worden toegepast. Jeugdigen die volgens het Preselect als midden of hoog risico voor recidive worden beschouwd, worden doorverwezen naar het Openbaar Ministerie (OM) of de Zittende Magistratuur (ZM), de rechter. In dat geval wordt door de RvdK een uitgebreider instrument, de Ritax A, afgenomen, waarmee tevens dynamische risicofactoren worden gemeten. Bij jeugdigen die volgens de Ritax A als midden- of hoog risico worden geclassificeerd, wordt een uitgebreider instrument, de Ritax B, ingezet. Dit instrument kijkt tevens naar de dynamische, beschermende factoren (Buysse & Petersen, 2023; Ministerie van JenV, 2024). Uiteindelijk worden de gegevens uit LIJ-afnames gebruikt om passende (strafrechtelijke) afdoeningen of interventies voor jeugdige justitiabelen te adviseren.

Zoals eerder benoemd in paragraaf 2.2, richt deze haalbaarheidsstudie zich exclusief op het grootste gedeelte van het LIJ, namelijk de Ritax A en Ritax B die worden afgenomen door de RvdK. Wanneer in de onderzoeksbevindingen hieronder wordt gesproken over het LIJ, wordt hiermee bedoeld op deze onderdelen, tenzij anders vermeld.

3.1.3 *Eerder onderzoek naar aanpassingen en de betrouwbaarheid van het LIJ*

Uit de literatuur blijkt dat het LIJ over de tijd aan verschillende ontwikkelingen en aanpassingen onderhevig is geweest. Er zijn diverse onderzoeken geweest naar de praktijkervaringen met en knelpunten omtrent instrumenten van het LIJ die hebben bijgedragen aan deze ontwikkelingen. Zoals omschreven door Nauta en Loef (2010), startte de eerste pilot van het LIJ in 2007. In deze pilot werd het instrument voor het eerst in de praktijk getest in de politieregio's Utrecht en Rotterdam-Rijnmond. Hoewel het gebruik van het LIJ werd uitgerold op basis van landelijk vastgestelde richtlijnen voor de werkwijze omtrent LIJ-afnames (Ministerie van JenV, 2011), bleek tijdens deze pilotfase dat die richtlijnen niet consistent werden nageleefd. Problematische inconsistenties werden onder andere vastgesteld bij de tijdigheid van het registreren van Ritax B-afnames. Dit was bijvoorbeeld problematisch voor de Jeugdreclassering, omdat zij binnen zes weken een plan van aanpak moet opstellen en daarvoor de informatie uit de Ritax B nodig heeft, die niet altijd op tijd beschikbaar was. In de regio Rotterdam-Rijnmond ontstonden ook problemen rondom het (tijdig) aanleveren van het kerndeel van LIJ-afnames in vroeghulp zaken, wat de betrouwbaarheid van LIJ-rapportages in deze pilotregio leek te beïnvloeden. Een gebrek aan tijdigheid kan de inhoudelijke betrouwbaarheid van LIJ-rapportages beïnvloeden, omdat op het moment dat beslissingen moeten worden genomen over de voortzetting van de zaak, welke hulp geboden moet worden, en welke straf of interventie het meest passend zou zijn, belangrijke informatie daartoe vanuit het LIJ niet beschikbaar kan zijn. Tevens waren er onduidelijkheden in de verdeling van verantwoordelijkheden tussen ketenpartners onderling. Andere noemenswaardige knelpunten kwamen naar voren bij LIJ-afnames in civiele zaken, waar (nog) niet structureel een LIJ werd afgenomen, hoewel dit volgens praktijkmedewerkers wel wenselijk zou zijn (Nauta & Loef, 2010). Daarnaast werd vastgesteld dat bij in verzekeringstelling (IVS)-zaken het LIJ niet consequent werd afgenomen, vooral in zaken waarbij jongeren in het weekend in verzekering werden gesteld.

In een vervolgstudie van Nauta en Loef (2011) werd de pilot uitgebreid naar meerdere regio's. Deze vervolgevaluatie benadrukte dat onduidelijkheden rondom taaktoewijzing en de implementatie van werkprocessen tussen ketenpartners nog steeds bestonden.

Zo voerde de RvdK soms Ritax-B afnames uit in zaken waarvoor de verantwoordelijkheid bij de Jeugdreclassering lag. Daarnaast bleven problemen bestaan omtrent LIJ-afnames in IVS-zaken, de samenwerking met civiele instanties, de toegankelijkheid van het LIJ-systeem onder ketenpartners, en bleek het rapportageformat moeilijk bruikbaar voor de rechterlijke macht (Nauta & Loef, 2011; Van der Put et al., 2011). Omdat de landelijke werkwijze, de validiteit en de betrouwbaarheid van het LIJ nog niet (uitgebreid) getoetst waren, ontbrak tevens voldoende draagvlak voor het instrument onder medewerkers in de praktijk (Nauta & Loef, 2011). Dit heeft mogelijk gevolgen gehad voor de consistentie van de registratie LIJ-data uit deze periode.

Alvorens de landelijke implementatie van het LIJ, werd de betrouwbaarheid van het LIJ gedeeltelijk onderzocht door Timmermans en Witvliet (2011). Hiervoor werden Ritax-gegevens van de RvdK gebruikt. Specifiek richtte dit onderzoek zich op de inhoudsvaliditeit³ van het domein 'psychisch disfunctioneren' en de interbeoordelaarsbetrouwbaarheid⁴ (IBB). Hieruit bleek dat de selectie van indicatoren voor psychische problematiek onvolledig was, waardoor belangrijke risicofactoren onvoldoende werden meegenomen in het meten van zorgbehoeften. Dit wees op beperkingen in de inhoudsvaliditeit van de Ritax bij het meten van één van de domeinen. Verder werd de betrouwbaarheid van het LIJ in deze studie niet uitgebreid onderzocht, en alhoewel de IBB als goed werd beoordeeld, was de steekproef in dit onderzoek te klein om definitieve conclusies te trekken over de IBB van de Ritax (of het LIJ als geheel) (Timmermans & Witvliet, 2011).

Enkele jaren later werd de IBB van het LIJ opnieuw en uitgebreider onderzocht door Van der Ark et al. (2018). Ook in dit onderzoek lag de focus expliciet op (gegevens uit) de Ritax van de RvdK. Ondanks methodologische uitdagingen, wat opnieuw leidde tot een beperkte steekproef ($N=61$, beoogd $N=150$), kwamen belangrijke bevindingen naar voren. Een kwetsbaarheid voor de betrouwbaarheid van de Ritax bleek de tijdsduur tussen het interview en het online invoeren van gegevens, die soms langer dan een week was. Dit vertraagde invoerproces kan leiden tot onnauwkeurigheden, omdat raadsonderzoekers afhankelijk zijn van hun herinnering aan het interview. Daarnaast kunnen details of nuances die tijdens het interview naar voren kwamen, verloren gaan of vermengd raken met informatie uit andere dossiers, wat de betrouwbaarheid van de ingevoerde gegevens verder zou kunnen beïnvloeden. Verder werd een aanzienlijk deel van de Ritax A en Ritax B (ongeveer 20%), als onduidelijk geformuleerd bevonden (Van der Ark et al., 2018). De algehele IBB bleek laag, wat overeenkomt met eerdere bevindingen over de lage IBB en inconsistenties in afnameprocedures tussen regio's (Nauta & Loef, 2010; Nauta & Loef, 2011; Timmermans & Witvliet, 2011). Uit de literatuur blijkt dus dat (regionale) verschillen in afnameprocedures een potentieel knelpunt vormen voor de betrouwbaarheid van Ritax-data over de tijd. Dit speelt mogelijk nog steeds, zoals onderstreept door interne documentatie, waaruit blijkt dat het landelijke gemiddelde van Ritax-afnames 80% bedraagt, maar regionaal sterk uiteen kan lopen, van tijdelijk 45% in de regio Amsterdam tot tijdelijk 90% in de regio Alkmaar (RvdK, 2018; RvdK, 2024). Dit terwijl een LIJ-afname het uitgangspunt is in alle strafzaken, als dat mogelijk is.⁵ Een studie naar risicoprofielen van jeugdige verdachten met Ritax-data uit 2012 en 2013

³ In hoeverre een specifiek concept adequaat gemeten wordt.

⁴ De mate van overeenstemming in het scoren van gedrag tussen beoordelaars.

⁵ Er zijn legitieme redenen voor de RvdK om af te wijken van het uitgangspunt 'LIJ-gebruik, tenzij' namelijk als het LIJ onbetrouwbaar wordt door onvoldoende antwoorden op de te stellen vragen (redenen: jongere weigert of is niet in staat tot medewerking, Alleenstaande Minderjarige Vreemdelingen die onvoldoende inbedden in Nederland, enz.).

benadrukte eveneens dat een deel van jeugdige strafzaken niet is opgenomen in Ritax-data, ondanks het uitgangspunt Ritax/LIJ, tenzij (Mensink et al., 2021). Dit roept vragen op over de dekking van Ritax-data, zowel tussen verschillende regio's als voor diverse typen strafzaken.

Verder is recentelijk ook de constructvaliditeit⁶ van het LIJ, aan de hand van Ritax A en B gegevens voor recidivevoorspelling beoordeeld in een normeringsonderzoek door Van der Put en Kapteijns (2020). Zij constateerden dat bijna de helft van de items binnen de tien domeinen van de Ritax niet voldeed aan de minimumcriteria voor samenhang met recidive, vooral in de domeinen werk, alcohol/drugs en geestelijke gezondheid. De bevindingen van dit normeringsonderzoek zijn verwerkt in de grootste aanpassing van de Ritax sinds de invoering ervan, ook wel bekend geworden als de ontwikkeling van de 'Ritax 2.0' (DGSenB, 2022; Buysse & Petersen, 2023; Van der Put & Kapteijns, 2020). Significante aanpassingen aan de nieuwe Ritax 2.0 waren, onder andere, de reductie van tien naar negen domeinen en het verwijderen van items met onvoldoende voorspellende waarde. Daarnaast zijn zes nieuwe items toegevoegd over online criminaliteit, en werd de categorische indeling van de totale risicoscore, het DRP, verfijnd. Voorheen kon deze worden opgedeeld in laag, midden of hoog, en sinds de Ritax 2.0 in heel laag, laag, midden, hoog en heel hoog (Buysse & Petersen, 2023; DGSenB, 2024a, 2024b; Ministerie van JenV, 2022). Overige aanpassingen betroffen het versimpelen van vraagstellingen en antwoordcategorieën. Ook werd de differentiatie naar leeftijd en sekse verwijderd, aangezien deze niet samenhangen met recidive (Van der Put & Kapteijns, 2020; Van der Put & Stolwijk, 2022).

3.1.4 *Ongewijzigde items over de tijd op basis van de grijze literatuur*

Bovengenoemde en andere aanpassingen in geregistreerde Ritax-gegevens als gevolg van wijzigingen aan het LIJ over de tijd, worden sinds 2019 systematisch bijgehouden in een logboek door dataspecialisten van de afdeling Ketenregie van het DGSenB binnen het ministerie van JenV. Uit dit logboek blijkt dat de aanpassingen aan het LIJ sinds de implementatie van de Ritax 2.0 vooral betrekking hebben op gewijzigde vraagstellingen en het toevoegen of verwijderen van antwoordcategorieën (DGSenB, 2024a). Sinds 2024 bestaat het LIJ (de Ritax 2.0) uit 108 items, verdeeld over 9 domeinen (Ministerie van JenV, 2024). In het logboek is te zien dat door de implementatie van de Ritax 2.0 in 2022 de meeste items zijn aangepast. Om vast te stellen welke items over de tijd heen ongewijzigd zijn gebleven, kijken we naar de items die in 2022 niet zijn aangepast. Vervolgens blikken we terug op de jaren 2019, 2020 en 2021, en vooruit naar 2023 en 2024 in het logboek om te beoordelen of deze in 2022 onveranderde items in de jaren daarvoor en daarna hetzelfde zijn gebleven. Uit het logboek blijkt dat een aantal items in de jaren voor 2022 in de loop der tijd zijn gewijzigd. Bijvoorbeeld: itemnummer 1.3a (aanwezigheid van ouderlijk toezicht, passend bij de leeftijd) en itemnummer 3.2 (is er bij deze jongere sprake van een problematische vrijetijdsbesteding) bleven onaangepast in 2022, maar werden beide in 2019 aangepast, waarbij een onderscheid werd gemaakt tussen online en offline ouderlijk toezicht. Sommige items zijn in de jaren daarna aangepast. Bijvoorbeeld, bij itemnummer 3.4 (in hoeverre is de jongere in staat om een baan te behouden) werd de antwoordcategorie 'weet niet wat nodig is om een baan te behouden' gewijzigd in 'heeft onvoldoende kennis, vaardigheden en/of motivatie om een baan te behouden'. Desondanks blijven er 50 items over die gedurende de periode 2019 t/m 2024 niet zijn

⁶ In hoeverre het instrument meet wat het moet meten.

aangepast, volgens het logboek (DGSenB, 2024a). Deze items zijn weergegeven in tabel 3.1.

Tabel 3.1 Ongewijzigde items over de tijd (2019 t/m 2024)

Domein en itemnummer	Itemnaam
Gezin	
1.2	Pleeggezin- en/of tehuisverleden
1.4a	Ouder(s)/verzorger(s) stelt/stellen regels op
1.4b	Ouder(s)/verzorger(s) ziet/zien toe of de jongere zich aan de regels houdt
1.5	Gehoorzaamheid van de jongere aan ouders/verzorgers (de mate waarin de jongere zich aan afspraken en regels houdt)
1.9	Is de jongere wel eens door een GI (Gecertificeerde Instelling) voor JB en JR vrijwillig of een kinderrechter civielrechtelijk uit huis geplaatst?
1.10	Is de jongere wel eens van huis weggelopen of uit huis gezet?
1.16	Financiële situatie van het gezin waar de jongere nu woont
1.19	Ernst/heftigheid van conflicten tussen gezinsleden
School	
2.1	Aantal dagen dagbesteding (school, stage en/of werk) per week.
2.4a	Schoolsituatie jongere afgelopen 6 maanden
2.5	De jongere hecht belang aan onderwijs/opleiding
2.8	Schoolprestaties van de jongere gedurende de afgelopen 6 maanden
2.9a	Gedrag van de jongere op school in afgelopen 6 maanden
2.9b	Kruis aan welke actie heeft plaatsgevonden
2.10	Spijbelt de jongere?
2.12a	Aantal schorsingen in de afgelopen 6 maanden
2.12b	Leeftijd bij de eerste schorsing in zijn/haar leven
2.13	Inschatting van de school dat de jongere de school/opleiding positief zal afronden
Werk, vrije tijd, financiën	
3.5	Huidige interesse in een baan
3.6	Oordeel jongere over de eigen financiële situatie
Relaties	
4.2	Bewondering voor en/of imiteren van antisociaal gedrag van leeftijdgenoten
4.3	De mate waarin de jongere weerstand kan bieden aan de invloed van leeftijdgenoten met antisociaal gedrag
4.4	Huidige positieve relaties met volwassenen, anders dan gezin of volwassenen die verbonden zijn aan werk of school
4.5	Huidige sociale binding met de gemeenschap

Domein en itemnummer	Itemnaam
4.6	Is de jongere wel eens gepest?
4.7	Heeft de jongere wel eens gepest?
Geestelijke gezondheid	
6.1	Is er sprake van een verstandelijke beperking?
6.2a	Is er een diagnose van psychische problemen die nu nog geldig is?
6.2c	Is medicatie voorgeschreven?
6.3	Staat de jongere onder behandeling van een psycholoog/psychiater?
6.4	Stemming
6.5	Automutilatie
6.6.	Gevaarlijke dingen vroeger
6.7	Suïcidale gedachten of gedragingen
6.8a	Slachtoffer van verwaarlozing
6.11	Getuige van huiselijk geweld
6.12	Ervaringen met geweld buiten het gezin
6.13a	Seksueel misbruik
Attitude (houding)	
7.2	Respect voor eigendom van anderen
7.5	Belangrijkste motief voor delictpleging(en)
7.6	Denkfouten
7.7	Empathie, gewetenswroeging, sympathie of gevoel voor slachtoffer(s)
7.8	Problemen op het gebied van de morele ontwikkeling
7.9	De mate waarin de jongere vindt dat wetten en regels en de gangbare omgangsvormen in het maatschappelijk verkeer voor hem/haar gelden
7.11	De mate waarin jongere bereid is te voldoen aan (de voorwaarden van) een maatregel of interventie gericht op gedragsverandering
Agressie	
8.1	Mate waarin jongere vindt dat schreeuwen en verbale agressie geschikt zijn voor het oplossen van een conflict
8.2	Mate waarin jongere vindt dat vechten en fysieke agressie geschikt zijn voor het oplossen van een conflict
8.5	Beheerst/gebruikt alternatieven voor agressief gedrag
Vaardigheden	
9.2	Impulsiviteit, handelen alvorens na te denken
9.6	Denken in termen van oorzaak en gevolg

3.1.5 Tussenconclusie

Op basis van de literatuurstudie blijkt dat er door de jaren heen verschillende wijzigingen aan onderdelen van het LIJ en knelpunten in afname en registraties zijn

geweest. Deze veranderingen kunnen de vergelijkbaarheid van de LIJ-gegevens van de Ritaxen A en B – wat de focus van deze haalbaarheidsstudie betreft – over de tijd beperken. Zoals weergegeven in tabel 3.2 zijn er wijzigingen geweest aan de inhoud. Met de introductie van de Ritax 2.0 in 2022 is een grote, significante wijziging toegepast in dit deel van het LIJ. Het aantal domeinen is gereduceerd en items die onvoldoende voorspellende waarde hebben voor recidive zijn verwijderd. Daarnaast zijn enkele nieuwe items toegevoegd en zijn vraagstellingen en antwoordopties aangepast. Deze wijzigingen, zowel wat betreft domeinscores als vraagstellingen en antwoordcategorieën, maken dat gegevens van vóór februari 2022 mogelijk niet goed vergelijkbaar is met de data daarna. Deze methodebreuk roept vragen op over de geschiktheid van oudere 'Ritax 1.0' gegevens voor analyses van trends en ontwikkelingen in jeugdcriminaliteit.

Ondanks een landelijk gestandaardiseerde werkwijze zijn er ook aanwijzingen voor knelpunten in de werkwijze in de afname van de instrumenten in de praktijk die een rol kunnen spelen in de vergelijkbaarheid. Zoals (regionale) verschillen in werkwijze, ook wat betreft afnames in IVS- en civiele zaken. De IBB van instrumenten uit het LIJ is nog nooit volledig onderzocht, terwijl inconsistenties daarin de betrouwbaarheid van gegevens op groepsniveau significant kunnen beïnvloeden. Of er wijzigingen zijn geweest in het gebruikte ICT-systeem (het Intelligente Formulierenmodule (IFM) systeem) blijft onduidelijk op basis van de literatuur. Wél blijkt dat dit systeem in de beginfase van het LIJ niet toegankelijk was voor alle ketenpartners. Dit is in de loop der tijd verbeterd, maar suggereert dat gegevens uit de beginfase van het instrument mogelijk minder geschikt zou zijn voor trendanalyses.

Op basis van de literatuur constateren we dan ook dat de betrouwbaarheid van gegevens uit het LIJ om jaarlijkse ontwikkelingen op populatieniveau te onderzoeken mogelijk onvoldoende is. Wellicht zijn de gegevens vanaf Ritax 2.0 wel bruikbaar, mits hierin geen grote verschillen in afname (over de jaren, regionaal, e.d.). Echter, om dergelijke conclusies te bevestigen, is een gedetailleerder beeld van de bruikbaarheid van de gegevens nodig. Hierbij moet onderzocht worden of een soortgelijk beeld ontstaat op basis van andere bronnen, zoals de ervaringen van praktijkexperts met (aanpassingen aan instrumenten van) het LIJ en indicatoren in LIJ-datasets die inzicht geven in de invulling van geregistreerde data door de tijd heen. Dit onderdeel wordt in het empirische vervolg van dit onderzoek verder bestudeerd.

Tot slot, ondanks alle veranderingen in de data over de tijd zijn er in de jaren 2019 tot heden vijftig items niet veranderd. Mochten vergelijkingen over de tijd worden gemaakt, dan kan specifiek de betrouwbaarheid van registraties bij deze items worden onderzocht.

Tabel 3.2 Tussenconclusie op basis van de literatuurstudie

Zijn er inhoudelijke wijzigingen of knelpunten omtrent:	Bevinding
De inhoud van het LIJ?	Ja Ontwikkeling en implementatie van de Ritax 2.0 Mogelijk significante wijzigingen aan items binnen de datasets In jaren 2019-2024 zijn 50 items ongewijzigd gebleven
Gebruikte ICT-systemen?	Misschien Toegankelijkheid van geregistreerde LIJ-data in de loop der tijd verbeterd
Wijze van afname?	Ja Regionale verschillen Verschillen in afname omtrent IVS-zaken
Overige zaken?	Misschien Onzekerheden omtrent de IBB

3.2 Interviews en focusgroepen – betrouwbaarheidsvraag

Zoals beschreven in paragraaf 2.2.1 hebben twintig (praktijk)experts deelgenomen aan de interviews en focusgroepen ten behoeve van de betrouwbaarheidsvraag in deze haalbaarheidsstudie. De interviews met experts richtten zich op vier onderwerpen, namelijk ervaringen met het gebruik van LIJ-data, indrukken van de bruikbaarheid en betrouwbaarheid van LIJ-data, potentiële problemen met LIJ-data, en benodigde aanpassingen aan het LIJ of LIJ-data over de tijd. De focusgroepen met raadsonderzoekers concentreerden zich op drie onderwerpen namelijk, ervaringen met LIJ-afnames in de praktijk, de registratie van LIJ-afnames in onlinesystemen, en ervaring met veranderingen in registratiesystemen en LIJ-afname procedures over de tijd.

Op basis van deze kwalitatieve dataverzameling zijn verschillende onderzoeksresultaten naar voren gekomen, die onderverdeeld zijn in drie thema’s. Namelijk: 1) (in)consistentie in de afname en invulling van het LIJ, 2) veranderingen in de registratie van LIJ-afnames over de tijd, en 3) de bruikbaarheid en beperkingen van LIJ-data voor vergelijkingen over de tijd.

3.2.1 (In)consistentie in de afname en invulling

Een van de eerste thema’s die uit de interview- en focusgroepen naar voren komt, is dat er ook anno 2024 verschillen bestaan in hoe de Ritax wordt afgenomen en ingevuld door raadsonderzoekers. Over het algemeen geven zij aan goed getraind te zijn in de werkwijze voor dergelijke afnames. Toch blijkt dat ervaren raadsonderzoekers na verloop van tijd hun eigen werkwijze ontwikkelen, wat soms afwijkt van de gestandaardiseerde werkwijze. Hierdoor kunnen de werkwijzen van ervaren raadsonderzoekers verschillen van die van nieuwere medewerkers, die tijdens hun trainingen soms anders worden opgeleid. Dit geldt bijvoorbeeld voor de invulling van scores bij ontbrekende of onduidelijke informatie. Waar de ene raadsonderzoeker een onbekend item alsnog invult op basis van een algemeen beeld van de jongere, vult een ander ‘onbekend’ in voor elk item waarvan geen concrete informatie beschikbaar is. Daarnaast proberen sommige raadsonderzoekers onbekende items

systematisch aan te vullen op basis van beschikbare dossierinformatie, terwijl anderen informatie die niet direct beschikbaar is tijdens de afname als 'onbekend' markeren en dit zo laten. Bovendien werd in zowel de focusgroepen als de expertinterviews benoemd dat bepaalde vragen, zoals 'opleidingsniveau ouders' of 'detentieverleden van ouders', relatief vaak als 'onbekend' worden ingevuld. Medewerkers van DG SenB Ketenregie hebben in hun werkwijze een grens van 20% 'onbekend' vastgesteld als acceptabel, maar voor sommige items of domeinen ligt dit percentage structureel hoger aldus sommige respondenten. Dit suggereert dat er variatie bestaat in de mate van invulling. De volledigheid van beschikbare gegevens binnen een individuele afname kan dus variëren, afhankelijk van de werkwijze van zowel de betrokken raadsonderzoeker als de beschikbare informatie die daadwerkelijk kan worden verzameld tijdens het onderzoek.

Daarnaast werd in de interviews en focusgroepen opgemerkt dat er variatie is in de mate waarin de Ritax wordt afgenomen, wat betreft strafzaken waarvoor het LIJ verplicht is versus de gevallen waar het daadwerkelijk wordt ingezet. Het landelijk gemiddelde voor jeugdige strafzaken waarin een Ritax wordt afgenomen ligt rond de 80%, maar dit percentage fluctueert per regio (zie ook 3.1.3). Er zijn signalen dat in bepaalde regio's structureel minder frequent werd afgenomen in het verleden.

Sommige respondenten benoemden dat het lastig is om een onderzoek af te nemen in jeugdige strafzaken van complexere aard, zoals IVS-zaken, waarbij het moeilijk is om tijdig – voor vrijlating – een onderzoek af te nemen. Daarbij zijn er in de avonduren, weekenden en op feestdagen soms tekorten in personeelsinzet. Vergelijkbare moeilijkheden werden ook benoemd bij strafzaken met alleenstaande minderjarige vreemdelingen, waarbij taalbarrières en een beperkte sociale omgeving de afname bemoeilijken of zelfs onmogelijk maken. Deze factoren kunnen negatieve gevolgen hebben voor de bruikbaarheid en betrouwbaarheid van de Ritax-gegevens voor (trend)analyses, met name wat betreft de representativiteit van de populatie.

3.2.2 *Veranderingen in de registratie van afnames over de tijd*

Een tweede thema wat naar voren kwam tijdens de interviews en focusgroepen betreft de wijze waarop gegevens worden gescoord en geregistreerd in het IFM-systeem. Uit de literatuurstudie blijkt dat Ritax in de loop der jaren aan een grote aanpassing onderhevig geweest is, wat gevolgen kan hebben voor de vergelijking hedendaagse gegevens met data uit het verleden (zie ook 3.1.3). Deze grote wijziging, namelijk de overgang van de Ritax 1.0 naar de 'Ritax 2.0', werd herhaaldelijk benoemd door de respondenten.

Een verandering die verschillende keren benoemd werd, betrof de aanpassingen in de categorische indeling van de totaalscore van het algemeen risicoprofiel. Met de overgang van Ritax 1.0 naar Ritax 2.0 is dit veranderd van een 3-puntenschaal (laag, midden, hoog) naar een 5-puntenschaal (heel laag, laag, midden, hoog, en heel hoog). Meer specifieke, inhoudelijke wijzigingen sloegen op het verwijderen of toevoegen van bepaalde items. Zo werd benoemd dat vragen over 'werk' zijn verwijderd omdat deze volgens het normeringsonderzoek minder relevant zouden zijn voor recidive. Daarentegen zijn er recentelijk ook nieuwe modules toegevoegd over online criminaliteit en wapenbezit, om het instrument beter te doen aansluiten op de veranderende problematiek onder jongeren. Eén respondent benoemde tevens dat er recentelijk speciale formats zijn ontwikkeld voor bepaalde complexe strafzaken, zoals in het geval van alleenstaande minderjarige vreemdelingen. Dit geldt ook voor IVS-

zaken met meerdere raadkamerzittingen, waarbij geen nieuwe informatie naar voren komt of het advies gelijk blijft aan die tijdens de eerste raadkamerzitting.

Verder kwam in de focusgroepen naar voren dat de uitgebreidheid van rapportages in de loop der tijd kan veranderen, hoewel dat mogelijk afhankelijk is van de (rechtbank) regio. Waar rapportages vroeger in sommige regio's veel uitgebreider waren, ligt nu de nadruk op beknopter en gerichter rapporteren, in lijn met de wensen van de regionale rechtbank. In andere regio's kan dit juist precies andersom zijn. De uitgebreidheid of juist beknoptheid van rapportages kan dus per regio verschillen, afhankelijk van de wensen van andere partners in de jeugdstrafrechtketen. Verder werd opgemerkt dat het IFM-systeem waarin de gegevens worden ingevuld, technisch niet altijd even stabiel is. Vooral tijdens spoedzaken kan het systeem op onregelmatige tijden uitvallen, wat het registratieproces verstoort.

3.2.3 Bruikbaarheid en beperkingen voor vergelijkingen over de tijd

Hoewel het LIJ, en dan met name de Ritax-data, een waardevolle en rijke bron van informatie is, zijn er voor vergelijkingen over de tijd mogelijk enkele beperkingen aldus de respondenten. Zoals al werd beschreven in paragraaf 3.2.2, vormt de methodebreuk in het LIJ door de invoering van de Ritax 2.0 een mogelijke beperking voor het gebruik van Ritax-data voor dergelijke vergelijkingen. Met name data uit het verleden – voor de Ritax 2.0 – wordt hierdoor door sommige respondenten als moeilijk bruikbaar voor vergelijkingen over de tijd, en daardoor als verminderd betrouwbaar beschouwd.

Daarnaast werd benoemd dat onvolledig ingevulde dossiers en afgebroken onderzoeken kunnen leiden tot een incompleet beeld van bepaalde groepen jongeren. Echter, ook kwam naar voren dat een LIJ-afname alleen geregistreerd kan worden in het IFM-systeem als het 'gereed gemeld' wordt, wat alleen mogelijk is als de afname voldoende is ingevuld. Daarom werd benoemd dat het extra belangrijk is dat ontbrekende informatie niet een leeg veld blijft, maar als 'onbekend' wordt gescoord tijdens de registratie van de afname in het IFM-systeem. Pas wanneer een afname als voldoende compleet wordt beschouwd, kan deze ook daadwerkelijk gereed worden gemeld. Echter, daardoor is mogelijk onder sommige jongeren met een LIJ-afname veel ontbrekende, inhoudelijke informatie, ondanks dat een afname als volledig geregistreerd staat.

Een ander aspect betreft de dynamische aard van het instrument zelf. Zoals eerder naar voren kwam in de literatuurstudie op basis van het logboek (zie ook paragraaf 3.1.3) werd ook tijdens de interviews duidelijk dat het LIJ meerdere keren per jaar wordt aangepast op basis van signalen uit de praktijk. Dit betreft drie verschillende soorten aanpassingen, namelijk, het wijzigen van de vraagstelling van een item, het toevoegen of verwijderen van antwoordcategorieën binnen een item, en het toevoegen of verwijderen van daadwerkelijke items. Deze veranderingen kunnen implicaties hebben voor de vergelijkbaarheid van de LIJ-data over de tijd, aangezien wijzigingen in de (vraagstelling of antwoordcategorieën van de) items de consistentie van de meetresultaten kunnen beïnvloeden. Dit betekent dat het moeilijker wordt om trends en veranderingen over tijd te interpreteren, aangezien de informatie niet op dezelfde manier gemeten wordt over de tijd.

In overeenstemming met de literatuur kwam tijdens de interviews ook naar voren dat men zich ervan bewust is dat de IBB van (instrumenten uit) het LIJ onvoldoende is

onderzocht. Het blijft onduidelijk in hoeverre onderzoeken op een vergelijkbare wijze worden afgenomen door verschillende raadsonderzoekers. Verder werd benoemd dat in de praktijk vaak wordt afgeweken van het geautomatiseerde strafadvies, welke als slotonderdeel van het LIJ wordt gegenereerd. Deze afwijkingen vereisen een gemotiveerde toelichting van de raadsonderzoeker, maar vanwege AVG- en privacywetgeving zijn deze gegevens niet structureel beschikbaar voor onderzoeken waarin LIJ-data gebruikt wordt. Deze en de eerdergenoemde beperkingen kunnen de bruikbaarheid van gegevens uit het LIJ voor onderzoek naar ontwikkelingen over de tijd echter beperken.

3.2.4 *Tussenconclusie*

Uit de resultaten van de interviews en de focusgroepen blijken eveneens verschillende factoren die de vergelijkbaarheid van gegevens uit het LIJ over de tijd beperken. Hoewel het LIJ als instrument landelijk is gestandaardiseerd, zijn er in de praktijk variaties in de wijze waarop raadsonderzoekers individuele onderzoeken afnemen en invullen. Dit komt onder andere door verschillen in werkervaring, interpretatie van de weging van scores en werkwijzen bij het omgaan met ontbrekende of onduidelijk informatie. Het is belangrijk hiermee rekening te houden als gegevens uit het LIJ worden gebruikt voor vergelijkingen over de tijd, omdat dergelijke verschillen in invulling mogelijk niet constant zijn over de tijd of regio's. Wanneer de mate of richting van deze variatie verandert, kunnen waargenomen ontwikkelingen in de data deels het gevolg zijn van wisselende invulpraktijken, in plaats van daadwerkelijke veranderingen in de populatie.

Verder heeft de overgang naar de Ritax 2.0 geleid tot significante wijzigingen in, onder andere, de totaalscore van het dynamisch risicoprofiel en de aard en omvang van het aantal items en domeinen.

Een ander aandachtspunt is informatie die mogelijk systematisch ontbreekt. Er zijn aanwijzingen dat sommige vragen relatief vaak als 'onbekend' worden gescoord. Dit kan gevolgen hebben voor de volledigheid van de data. Het is onduidelijk in welke mate, voor welke items en hoe vaak ontbrekende waarden structureel voorkomen. Daarnaast wordt mogelijk in sommige regio's en onder bepaalde soorten strafzaken het LIJ minder consistent afgenomen. Dit kan gevolgen hebben voor de representativiteit van bepaalde groepen binnen de gehele populatie van jeugdigen waarbij een instrument uit het LIJ is afgenomen. Echter op basis van de interview- en focusgroep data blijft onduidelijk in welke mate dit zo is.

De interviews en focusgroepen met experts en ervaringsdeskundigen in het gebruik van het LIJ bevestigen daarmee de bevindingen uit de literatuurstudie. Variatie in afname en invoer van gegevens tussen medewerkers en regio's, inhoudelijke veranderingen aan het instrument over de tijd en mogelijke variatie in het veelvuldig registreren van informatie als onbekend maken dat er knelpunten zijn voor het gebruik van de gegevens voor vergelijkingen over de tijd. In hiernavolgende empirische analyses wordt op basis van recente gegevens uit de Ritax kwantitatief nagegaan welke mogelijke bias er is in de registratie van ontbrekende informatie (paragraaf 3.3). Dit biedt verdere inzichten in de volledigheid, en derhalve, de bruikbaarheid van de data voor het analyseren van ontwikkelingen over de tijd.

Tabel 3.3 Tussenconclusie op basis van de interviews en focusgroepen

Zijn er signalen uit de interviews en focusgroepen welke duiden op:	Bevinding
Significante veranderingen (in afname-instructies) over de tijd?	Ja Overgang naar 'Ritax 2.0' Periodieke aanpassingen aan of binnen de items
Ongewone regionale variatie?	Misschien Variaties in werkwijzen omtrent LIJ-afnames en invulling onder raadsonderzoekers en regio, niet duidelijk of dit een vast patroon is
Systematisch ontbrekende gegevens?	Misschien Op sommige vragen wordt structureel veel 'onbekend' gescoord, nog geen sluitend bewijs op welke schaal dit plaatsvindt

3.3 Ontbrekende waarden analyse – Betrouwbaarheidsvraag

In deze paragraaf worden de kwantitatieve resultaten van de betrouwbaarheidsvraag gepresenteerd. Allereerst wordt een overzicht gegeven van enkele beschrijvende kenmerken van de jongeren. Zoals benoemd in paragraaf 1.1.2 richtten de kwantitatieve analyses zich voornamelijk op de betrouwbaarheid van de data in termen van de mate van ontbrekende waarden. Daarbij werd onderzocht de mate waarin de verschillende variabelen (items) zijn ingevuld over de hele groep in de onderzochte datasets. Deze kwantitatieve analyse richt zich specifiek op de mate waarin gegevens op item niveau per jaar al dan niet zijn ingevuld. Er zijn verschillende datasets verzameld op itemniveau over meerdere jaren. Vervolgens is gekeken naar de mate van ontbrekende gegevens in de verschillende (domeinen en) items. Daarnaast is onderzocht in hoeverre het aantal ontbrekende waarden op items verschilt tussen subgroepen zoals geslacht, leeftijd, afnameregio en instrumenttype. Delict zwaarte is hierbij niet meegenomen, aangezien uit de literatuur, interviews en focusgroepen geen aanwijzingen naar voren kwamen dat de volledigheid van de data zou verschillen naar delict zwaarte.

Er zijn vier verschillende datasets gebruikt met gegevens op itemniveau. Die gezamenlijk de jaren 2020 tot en met 2023 omvatten. De data zijn geleverd door afdeling ketenregie van DGSenB.

De ontbrekende waarden analyse begon met de dataset van 2023. Hieruit bleek vrij snel dat er regelmatig ontbrekende waarden voorkomen op verschillende items. Bovendien toonde het kwalitatieve gedeelte van dit onderzoek al aan dat met de introductie van de Ritax 2.0 een significante methodebreuk is ontstaan, waardoor de data van voor 2022 mogelijk niet langer geschikt is voor vergelijkingen over de tijd. Om deze redenen is ervoor gekozen om de data op itemniveau uitsluitend te analyseren voor de jaren 2022 en 2023.

3.3.1 Data verkennen en definiëren van ontbrekende waarden

Alvorens het uitvoeren van de analyses op ontbrekende waarden, zijn de aangeleverde databestanden verkend. Dit betrof onder andere, het genereren van beschrijvende

statistieken, het vaststellen van welke items tot welke domeinen behoorden, en het ontdebellen van afnames in gevallen waarin meerdere afnames aanwezig waren binnen hetzelfde zaaknummer voor één en dezelfde jongere. Na overleg met een LIJ-superuser is voor het ontdebellen telkens de meest recente versie (op basis van het hoogste versienummer) geselecteerd, met een voorkeur voor een Ritax B indien aanwezig. Dit ontdebellen was noodzakelijk om een vertekend beeld in het aantal ontbrekende waarden te voorkomen.

Zoals benoemd in paragraaf 2.1 is gekeken naar de ontbrekende waarden op elke variabele (item) voor de hele groep. Inhoudelijk kon dit onderverdeeld worden in twee soorten ontbrekende waarden, namelijk 'onbekend' scores of lege cellen op itemniveau. Binnen deze twee typen ontbrekende waarden kon onderscheid gemaakt worden tussen verklaarbare en onverklaarbare ontbrekende waarden.

In de handleiding van het LIJ van de RvdK (Ministerie van JenV, 2024⁷) is te lezen dat bepaalde items afhankelijk zijn van een score op een eerder item: ze zijn conditioneel. Dit levert verklaarbare ontbrekende waarden op. Verklaarbare ontbrekende waarden zijn items die op basis van de vraagstructuur logischerwijs leeg moesten blijven. Bijvoorbeeld, als een jongere niet binnen een gezin woonachtig is, dan zijn een aantal items binnen het domein gezin, zoals de mate van ouderlijk toezicht en of dat er thuis regels worden opgesteld, niet van toepassing. Verschillende items in het LIJ-instrument zijn dus alleen van toepassing indien op een eerder item een specifieke waarde is ingevuld (Ministerie van JenV, 2024). Sommige items zijn daarnaast afhankelijk van het type afname, zoals een volwaardige Ritax A- of Ritax B-afname, of een verkorte IVS2A-afname. De lege velden op deze items vallen dan onder verklaarbare ontbrekende waarden.

In het kader van dit onderzoek was het belangrijk dat verklaarbare ontbrekende waarden niet ten onrechte als onverklaarbare ontbrekende waarden werden geïnterpreteerd. Voor dit doeleinde is een uitgebreide set van conditionele regels toegepast op de datasets. Items die niet van toepassing waren vanwege een conditionele afhankelijkheid zijn gehercodeerd en daarmee onderscheiden van onverklaarbare ontbrekende waarden.

De ontbrekende waarden die vervolgens overbleven waren 1) lege cellen op itemniveau die wél ingevuld hadden moeten worden, of 2) velden die expliciet als 'onbekend' waren gemarkeerd, waarbij er dus wel een aanwijzing was dat er informatie was, maar de inhoudelijke informatie ontbrak waardoor het item niet ingevuld kon worden en als 'onbekend' is geregistreerd.

Vervolgens is binnen elk jaarcohort voor alle items over de kolommen geanalyseerd waar en hoe vaak dergelijke 'onverklaarbare' ontbrekende waarden voorkwamen. De verwachting was dat er weinig tot geen lege cellen op itemniveau gevonden zouden worden gevonden. Dit omdat de datasets uitsluitend gereed gemelde afnames bevatten, wat inhoudt dat items voldoende ingevuld moeten zijn op individueel niveau.

Zoals hierboven toegelicht, was bovendien gecorrigeerd voor verklaarbare ontbrekende in het kader van conditioneel afhankelijke items. Aan de andere kant kan bij vrijwel elk item een 'onbekend' score worden toegekend. Wat betreft deze 'onbekend' scores was de verwachting dat meerdere items een opvallend hoog aantal

⁷ Zie ook: <https://www.rijksoverheid.nl/documenten/richtlijnen/2025/01/15/handleiding-landelijk-instrumentarium-jeugdstrafrechtken-lij>

van deze scores zouden vertonen (door meerdere respondenten aangegeven, met name door onderzoekers die in het verleden met LIJ-data hebben gewerkt).

Binnen de datasets op itemniveau is met een chi-kwadraattoets getoetst of onverklaarbare ontbrekende waarden structureel vaker voorkwamen binnen bepaalde subgroepen, op basis van enkele beschikbare kenmerken in de datasets. Namelijk, geslacht, leeftijd, afnameregio, delictzwaarte en instrumenttype. Gezien het grote aantal toetsen en het bijbehorende risico op kanskapitalisatie is hierbij een strenge grenswaarde gehanteerd van $p < 0,01$.

De resultaten bespreken we hieronder per jaar, beginnend met 2023. In de tabellen en figuren worden de jaren 2023 en 2022 naast elkaar gepresenteerd.

3.3.2 *Achtergrondkenmerken*

Zoals in de inleiding benoemd, is in deze haalbaarheidsstudie gekeken naar een aantal achtergrond kenmerken om iets te kunnen vernemen over de populatie (seks, leeftijd en afnameregio) en het delict type (delict zwaarte). In tabel 3.4 wordt de procentuele verdeling op deze kenmerken weergegeven voor de jaren 2022 en 2023. De tabel geeft een overzicht van de verdeling naar leeftijd, geslacht, delict zwaarte en de regio van afname.

Voor het jaar 2023 was de gemiddelde leeftijd van jongeren in de dataset 15,4 jaar ($SD=1,3$). Deze leeftijdsverdeling komt vrijwel overeen met die van 2022, waarin de gemiddelde leeftijd eveneens 15,4 jaar was, hoewel in het 2023 dataset enkele ($N=5$, 0,1%) 18-jarige jongeren zaten. Wat betreft geslacht is er in beide jaren een duidelijke meerderheid van jongens. In 2023 ging het om 88,6% jongens en 11,4% meisjes, tegenover 90,2% jongens en 9,8% meisjes in 2022. Bij delict zwaarte valt op dat het aandeel middelzware en zware delicten in 2023 iets hoger is dan in 2022. In 2023 werd 36,5% van de delicten als middelzwaar en 31,8% als zwaar gecategoriseerd, terwijl deze percentages in 2022 iets lager (34,6% versus 31,1%) waren. Het aandeel lichte delicten was in 2023 iets lager (20,6%) dan in 2022 (21,5%). Voor de afnameregio's komen in beide jaren dezelfde drie regio's het vaakst voor. Dit waren Rotterdam (16,1%, 2023 - 16,2%, 2022), Den Haag (12,1%, 2023 - 12,9%, 2022), en Arnhem (9%, 2023 - 8,9%, 2022).

Samengevat laten de beschrijvende statistieken zien dat de jongeren in de dataset van 2023 vergelijkbaar zijn met die van 2022 wat betreft leeftijd, geslacht en afnameregio. Hoewel er iets hogere percentages van middelzware en zware delicttypen voor het jaar 2023 voorkomen dan in 2022, is het percentage onbekend in beide jaren relatief hoog. Daarom kunnen we weinig zeggen over verschillen in zwaarte van het delict tussen beide jaren.

Tabel 3.4 Achtergrondkenmerken jeugdigen met een Ritax A of Ritax B afname in 2022 en 2023

Variabele	2022 (N, %) of M (SD)	2023 (N, %) of M (SD)
Leeftijd		
Gemiddelde	15,4 (1,3)	15,4 (1,3)
Min/max	12-17	12-18
Geslacht		
Man	4350 (90,2%)	4865 (88,6%)
Vrouw	471 (9,8%)	629 (11,4%)
Delictzwaarte		
Licht	1035 (21,5%)	1134 (20,6%)
Middel	1670 (34,6%)	2003 (36,5%)
Zwaar	1499 (31,1%)	1746 (31,8%)
Onbekend	617 (12,8%)	611 (11,1%)
Afnameregio		
's-Hertogenbosch	140 (2,9%)	145 (2,6%)
Alkmaar	165 (3,4%)	217 (3,9%)
Almelo	123 (2,6%)	162 (2,9%)
Amsterdam	295 (6,1%)	322 (5,9%)
Arnhem	430 (8,9%)	496 (9,0%)
Breda	319 (6,6%)	364 (6,6%)
Den Haag	621 (12,9%)	664 (12,1%)
Eindhoven	188 (3,9%)	223 (4,1%)
Groningen	204 (4,2%)	286 (5,2%)
Haarlem	139 (2,9%)	240 (4,4%)
Leeuwarden	225 (4,7%)	223 (4,1%)
Lelystad	214 (4,4%)	274 (5,0%)
Maastricht	172 (3,6%)	173 (3,1%)
Middelburg	131 (2,7%)	137 (2,5%)
Roermond	174 (3,6%)	162 (2,9%)
Rotterdam	780 (16,2%)	885 (16,1%)
Utrecht	357 (7,4%)	361 (6,6%)
Zwolle	138 (2,9%)	151 (2,7%)

3.3.3 Ontbrekende waarden op itemniveau

Zoals hierboven benoemd is gekeken naar twee soorten onverklaarbare ontbrekende waarden, namelijk 1) daadwerkelijk lege cellen op itemniveau en 2) 'onbekend' scores. De conditionele, oftewel, verklaarbare ontbrekende waarden waarvoor gecorrigeerd is

zijn dus geen onderdeel van de analyses, en worden volledig buiten beschouwing gelaten in hieronder genoemde aantallen.

Ontbrekende waarden op itemniveau onder het gehele cohort

Wat betreft lege cellen op itemniveau bleek dat het percentage onverklaarbare lege cellen op itemniveau in de dataset voor het jaar 2023 laag was. Binnen het domein 'gezin' werd 0,1% aan lege velden gevonden, bij het item 'opleidingsniveau moeder'. Binnen het domein 'school' waren meerdere items met een klein percentage lege velden, variërend van 0,1% bij het item 'schoolsituatie laatste opleiding' tot 2,7% bij de items 'verwijderd les afgelopen 6 maanden' en 'schooldiploma'. Binnen het domein 'geestelijke gezondheid' werd 2,9% lege velden gevonden bij het item 'gevaarlijke dingen vroeger'. In de overige domeinen werden geen onverklaarbare lege velden aangetroffen. Kortom, er bleken weinig items met onverklaarbare lege cellen te zijn. Dit hadden we verwacht aangezien een afname alleen geregistreerd wordt als alle velden zijn ingevuld⁸.

Uit de analyse van onverklaarbare, ontbrekende waarden geregistreerd die als 'onbekend' stond geregistreerd, ontstond een ander beeld. Meerdere items hadden meer dan 5%, 10% of 20% aan 'onbekend' scores (gerekend over de kolommen). Voor een volledig overzicht van de aantallen en percentages van ontbrekende waarden op alle items in het dataset van 2023 en 2022, zie bijlage 1. De top drie items met het hoogste percentage 'onbekend'-scores waren hetzelfde voor beide jaren:

- Domein gezin — 'opleidingsniveau vader' (44,6% — 2023, 44% — 2022)
- Domein gezin — 'opleidingsniveau moeder' (37,2% — 2023, 38,6% — 2022)
- Domein gezin — 'beroepsniveau vader' (34,3% — 2023, 33,6% — 2022)

De top drie items met het laagste percentage 'onbekend'-scores waren ook hetzelfde voor beide jaren:

- Domein agressie — 'seksueel gedrag' (0% — 2023, 0% — 2022)
- Domein agressie — 'gewelddadig gedrag' (0% — 2023, 0% — 2022)
- Domein school — 'gedrag afgelopen zes maanden-acties' (0% — 2023, 0% — 2022)

Tabel 3.5 geeft een overzicht van het aantal en percentage items dat afvalt per domein gezien de ontbrekende waarden op de items binnen deze domeinen, dit voor beide jaren. Het referentiejaar is 2023, omdat dit het meest recente jaar is waarvoor volledige data beschikbaar was voor deze haalbaarheidsstudie. Door dit jaar als uitgangspunt te nemen, kan eerst de actuele stand van zaken wat betreft de volledigheid van de invulling van de data worden beoordeeld. Vervolgens is een eerder jaar, namelijk 2022, als vergelijking gebruikt om veranderingen in invulling over de tijd inzichtelijk te maken. Verder wordt in tabel 3.5 onderscheid gemaakt tussen verschillende percentages van ontbrekende waarden. Zoals beschreven in de inleiding is specifiek gekeken naar het percentage ontbrekende waarden per item, en of dit binnen een cutoffwaarde van 5%, 10% of 20% viel. Deze percentages waren als volgt berekend. Voor ieder item zijn onverklaarbare lege cellen en alle 'onbekend'-scores als ontbrekende waarden gedefinieerd in het softwareprogramma. Vervolgens is aan de hand van frequentietabellen gekeken naar de aantallen en percentages ontbrekende waarden per item, over het hele jaar. Daarna is per item beoordeeld of het percentage ontbrekende waarden binnen een maximale drempelwaarde waarde van 5%, 10% of 20% viel. Aansluitend is berekend hoeveel items per domein zouden afvallen als de

⁸ De reden waarom een klein aantal daadwerkelijke, lege cellen voorkomt in de data is vooralsnog onbekend.

laagste en meest strikte drempelwaarde van 5%, een 'middel' drempelwaarde van 10%, of de hoogste, meest ruime drempelwaarde van 20% zou worden gehanteerd. Tot slot is bepaald hoeveel items zouden afvallen per domein. Hiervoor is een percentage berekend, waarbij het aantal 'afvallende' items afgezet werd tegen het totaal aantal items in dat domein. Per domein is dus berekend welk percentage van de items zou afvallen vanwege een te hoog percentage ontbrekende waarden. Bij de laagste, meest strikte drempelwaarde van maximaal 5% ontbrekende waarden vallen dus meer items uit dan bij de hoge, meest ruime drempelwaarde van maximaal 20% ontbrekende waarden. Hieruit bleek dat bij een strikte drempel van maximaal 5%, meer dan de helft van de items in meer dan de helft van de domeinen (vijf van de negen) afvalt. Met name de domeinen relaties (77,8%), attitude (72,7%) en gezin (61,3%) hebben een aanzienlijk hoog percentage afvallende items. Bij een drempel van 10% valt in twee domeinen iets minder dan de helft van de items af, en in drie domeinen ruim een kwart. Bij een ruime drempel van 20% blijven in ongeveer de helft van de domeinen (vijf van de negen) alle items behouden.

Tabel 3.5 Aantal en percentage items dat afvalt per domein voor het jaar 2023 in vergelijking met het jaar 2022

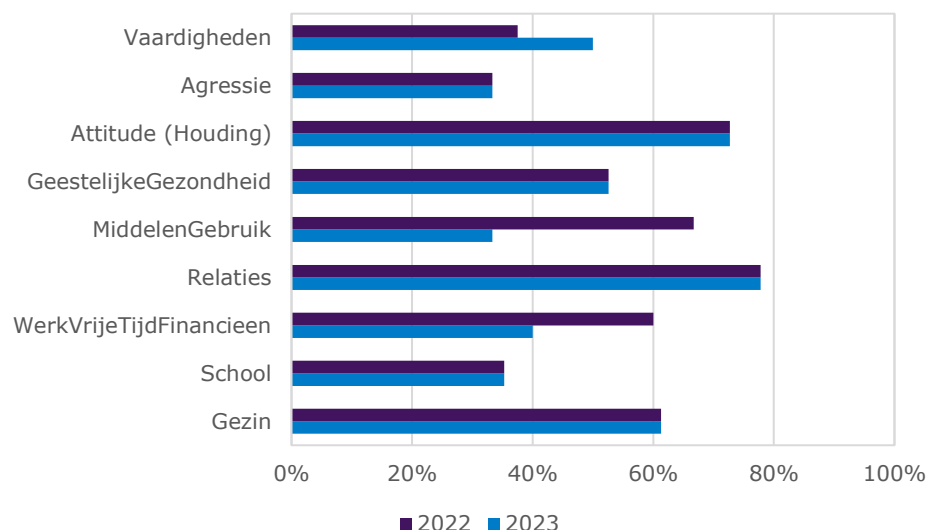
Domein	Totaal aantal items	Aantal afvallende items bij 5% cut-off	Percentage afvallende items bij 5% cut-off	2022 t.o.v. 2023*	Aantal afvallende items bij 10% cut-off	Percentage afvallende items bij 10% cut-off	2022 t.o.v. 2023	Aantal afvallende items bij 20% cut-off	Percentage afvallende items bij 20% cut-off	2022 t.o.v. 2023
Gezin	31	19	61,3%	=	14	45,2%	=	5	16,1%	+
School	17	6	35,3%	=	4	23,5%	-	0	0,0%	=
Werk, vrije tijd, financiën	10	4	40,0%	+	3	30,0%	+	2	20,0%	-
Relaties	9	7	77,8%	=	4	44,4%	=	3	33,3%	=
Middelen-gebruik	6	2	33,3%	+	0	0,0%	+	0	0,0%	=
Geestelijke gezondheid	19	10	52,6%	=	6	31,6%	-	2	10,5%	-
Attitude (houding)	11	8	72,7%	=	2	18,2%	=	0	0,0%	=
Agressie	6	2	33,3%	=	1	16,7%	+	0	0,0%	=
Vaardigheden	8	4	50,0%	-	0	0,0%	+	0	0,0%	=

* '+' is 2022 hoger dan 2023 (meer dan 5% verschil), '-' is 2022 lager dan 2023 (meer dan 5% verschil) en '=' is verschil kleiner dan 5%.

In de figuren 3.2, 3.3 en 3.4 is voor beide jaren visueel weergegeven hoeveel procent van de items afvalt, gezien de verschillende cut-offwaarden van maximaal 5% (figuur 3.2), 10% (figuur 3.3) en 20% (figuur 3.4) ontbrekende waarden. Naarmate het percentage ontbrekende waarden ruimer wordt genomen, vallen er steeds minder items af, en vice versa.

Zoals weergegeven in figuur 3.2 hebben de domeinen attitude en relaties hoge uitvalpercentages in beide jaren. Dit duidt erop dat de items binnen deze domeinen relatief vaak niet volledig zijn ingevuld. De domeinen agressie en school vertonen juist de laagste uitvalpercentages, wat erop wijst dat de items binnen deze domeinen relatief vaker volledig worden ingevuld.

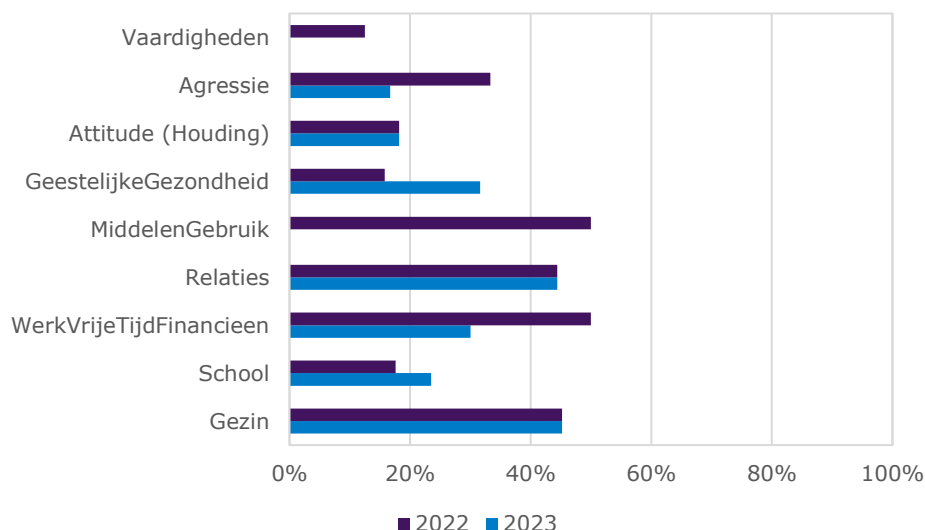
Figuur 3.2 Uitval van items 2022 vs. 2023 bij 5% cut-off



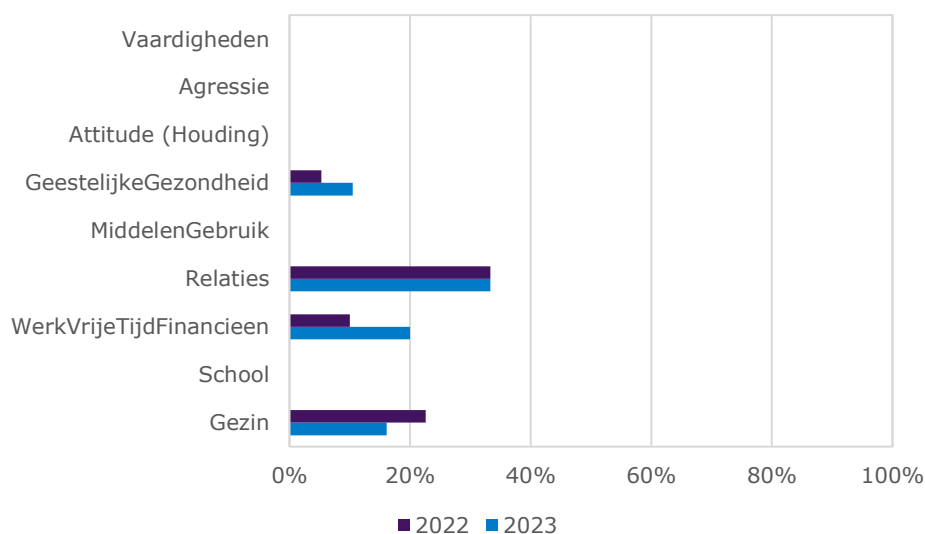
Bij een cut-offwaarde van 10% vallen aanzienlijk minder items af (zie figuur 3.3). Alleen in twee domeinen, namelijk gezondheid en relaties, blijft nog bijna de helft van de items afvallen. In de overige domeinen blijven de meeste items behouden, en lijken dus wel voldoende te zijn ingevuld gezien de iets ruimere cut-offwaarde van 10%. Zoals zichtbaar in figuur 3.5, blijven bij de meest ruime cut-offwaarde van 20% bij de meeste domeinen vrijwel alle items behouden. De mate van ontbrekende waarden in verschillende domeinen beperkt zijn tot kleine(re) percentages van maximaal 5 of 10%.

De verschillen tussen 2022 en 2023 zijn over het algemeen beperkt, ongeacht de cut-offwaarden of het type domein. De uitzonderingen hierop zijn de domeinen Werk, vrije tijd, financiën, middelengebruik en vaardigheden. Hier is in sommige cut-offcategorieën een iets groter verschil zichtbaar. De mate van ontbrekende waarden, oftewel, de mate van invulling van items varieert dus iets tussen de jaren, maar de algemene patronen van ontbrekende waarden blijft ongeveer consistent.

Figuur 3.3 Uitval van items 2022 vs. 2023 bij 10% cut-off



Figuur 3.4 Uitval van items 2022 vs. 2023 bij 20% cut-off



3.3.4 Ontbrekende waarden op itemniveau binnen verschillende subgroepen

Vervolgens is onderzocht of er binnen bepaalde subgroepen (geslacht, leeftijd), naar regio of type instrument verschillen zijn in de mate waarin er onverklaarbare ontbrekende waarden zijn. Met een significantiedrempel van $p < 0,01$ werden voor sommige items statistisch significante verschillen gevonden tussen subgroepen. Dit betekent dat bij deze items de kans op een ontbrekende waarde afhankelijk was van bijvoorbeeld, het geslacht of de leeftijd van de jongere wiens afname het betrof.

Geslacht

Bij de analyse van ontbrekende waarden op itemniveau is als eerste onderzocht of er een significant verband bestaat tussen het aantal ontbrekende waarden per item en

geslacht (jongen/meisje). Voor 2023 werd bij 19 van de 117 items een significant verschil naar geslacht gevonden. In 2022 betrof het 22 van de 117 items. De meeste van deze items waren binnen de domeinen gezin, school en attitude, waarbij jongens relatief vaker ontbrekende waarden hadden. Alleen binnen het domein geestelijke gezondheid waren er juist vaker ontbrekende waarden bij meisjes. Dit suggereert dat vooral bij jongens bepaalde thema's of onderwerpen vaker niet zijn ingevuld of geregistreerd. Tabel 3.6 toont de items met een significant tussen het aantal ontbrekende waarden per item en geslacht, evenals de richting van dat verband en de bijbehorende statistieken. Voor een volledig overzicht van alle items met een significant verband voor het jaar 2022 en 2023, zie bijlage 2.

Tabel 3.6 Items met significante verbanden tussen het aantal ontbrekende waarden per item en geslacht voor het jaar 2023

Item	Meer of minder missings bij jongens (+/-)	Statistiek (p -waarde, χ^2)
Praten Over Online Gedrag	+	$p=0,002$, $\chi^2=9,15$
Opleidingsniveau Vader	+	$p<0,001$, $\chi^2=23,1$
Opleidingsniveau Moeder	+	$p<0,001$, $\chi^2=35,89$
Beroepsniveau Moeder	+	$p=0,001$, $\chi^2=13,32$
Ondersteunend Netwerk	+	$p=0,001$, $\chi^2=10,61$
Conflicten Gezinsleden	+	$p<0,001$, $\chi^2=11,98$
Beloning Gewenst Gedrag	+	$p=0,002$, $\chi^2=9,61$
Spijbelen	+	$p=0,004$, $\chi^2=8,25$
Verwijderd Les Afgelopen 6 Maanden	+	$p=0,009$, $\chi^2=6,82$
Schorsing Afgelopen 6 Maanden	+	$p=0,003$, $\chi^2=9,05$
Positieve Relaties	+	$p<0,001$, $\chi^2=15,95$
Is Gepest	+	$p=0,003$, $\chi^2=8,73$
Mishandeling Seksueel Signalen	-	$p<0,001$, $\chi^2=106,58$
Motief Delict	+	$p<0,001$, $\chi^2=45,33$
Denkfouten	+	$p=0,008$, $\chi^2=7,01$
Empathie	+	$p<0,001$, $\chi^2=25,58$
Morele Ontwikkeling	+	$p=0,003$, $\chi^2=9,05$
Regels En Omgangsvormen	+	$p=0,010$, $\chi^2=6,71$
Alternatieven Agressie	+	$p=0,002$, $\chi^2=9,37$

Leeftijd

Daarnaast is ook onderzocht of er een significant en lineair verband bestond tussen het aantal ontbrekende waarden per item en leeftijd (twaalf tot en met achttien jaar). Voor 2023 werd bij 21 van de 117 items een significant, lineair verband gevonden tussen leeftijd en ontbrekende waarden. In 2022 ging het om 10 van de 117 items. In alle gevallen was dit een positief verband: hogere leeftijd ging samen met meer ontbrekende waarden. Dit patroon was het duidelijkst zichtbaar binnen de domeinen school en attitude. De sterkste verbanden werden gevonden bij school-gerelateerde

items, wat mogelijk wijst op moeilijkheden bij het invullen van items over het schoolverleden bij oudere jongeren. Tabel 3.7 toont de items met een significant verband tussen het aantal ontbrekende waarden per item en leeftijd, evenals de richting van dat verband en de bijbehorende statistieken. Voor een volledig overzicht van alle items met een significant verband voor het jaar 2022 en 2023, zie bijlage 2.

Tabel 3.7 Items met significante verbanden tussen het aantal ontbrekende waarden per item en leeftijd voor het jaar 2023

Item	Meer of minder missings bij hogere leeftijd (+/-)	Statistiek (p -waarde, gamma, sterkte)
Bestraffing Ongewenst Gedrag	+	$p=0,001$, 0,137, zwak
Problemen Vader Huidig Justitie	-	$p=0,046$, -0,069, zwak
Schoolsituatie	+	$p<0,001$, 0,364, matig
Schoolsituatie Laatste Opleiding	+	$p<0,001$, 0,423, matig
Schoolsituatie Reden Niet Op School	+	$p<0,001$, 0,390, matig
Speciaal Onderwijs Vanwege	+	$p<0,001$, 0,712, sterk
Gedrag Afgelopen 6 Maanden	+	$p=0,011$, 0,107, zwak
Leeftijd Eerste Schorsing	-	$p=0,010$, -0,231, matig
Online Tijdsbesteding	+	$p<0,001$, 0,115, zwak
Bewonderen Imiteren	+	$p<0,001$, 0,170, zwak
Weerstand Antisociaal	+	$p<0,001$, 0,185, zwak
Drugsgebruik Mate Van	+	$p=0,002$, 0,119, zwak
Gokgedrag	+	$p<0,001$, 0,155, zwak
Verstandelijke Beperking	-	$p=0,022$, -0,047, zwak
Accepteert Verantwoordelijkheid	+	$p=0,014$, 0,197, zwak
Kennis Van Digitale Wereld	+	$p=-$, 0,125, zwak
Motief Delict	+	$p<0,001$, 0,167, zwak
Empathie	+	$p=0,003$, 0,134, zwak
Toekomstbeeld	-	$p=0,039$, -0,061, zwak
Conflicten Oplossen Fysiek	+	$p=0,019$, 0,162, zwak
Impulsiviteit	+	$p=0,020$, 0,186, zwak

Afnameregio

Verder is gekeken naar significante verbanden tussen het aantal ontbrekende waarden en regio van afname. Voor beide jaren werd bij het overgrote deel van de items een significant verschil gevonden tussen ontbrekende waarden en de regio van afname. Oftewel, in sommige regio's worden afnames minder goed geregistreerd: in meerdere domeinen scoren bepaalde regio's consequent hoger op ontbrekende waarden. Dit suggereert mogelijke verschillen in werkwijze, registratiepraktijken of interpretatie van het instrument per locatie. Tijdens de interviews en focusgroepen kwam al naar voren

dat dergelijke verschillen inderdaad kunnen voortkomen uit variaties in de werkwijze en wijze van invulling tussen raadsonderzoekers. Deze kunnen regionaal, en soms zelfs per team verschillen.

In 2023 varieerde het percentage ontbrekende waarden per item naar regio van 0% tot 64,6%. De top drie regio's met het hoogste percentage aan ontbrekende waarden op een item in 2023 waren: Haarlem (64,6%, domein relaties – verschil contacten online en offline), Amsterdam (57,5%, domein gezin – opleidingsniveau vader), en 's-Hertogenbosch (42,1%, domein geestelijke gezondheid – verstandelijke beperking). In 2022 varieerde het percentage ontbrekende waarden per item naar regio van 0% tot 59%. De top drie regio's met het hoogste percentage aan ontbrekende waarden op een item in 2022 waren hetzelfde als in 2023, maar dan in andere volgorde. Namelijk, Haarlem (59%, domein gezin – opleidingsniveau vader), 's-Hertogenbosch (47,9%, domein geestelijke gezondheid – psychische problemen), en Amsterdam (42,7%, domein werk, vrije tijd, financiën).

Tabel 3.8 toont de top drie items met het hoogste percentage ontbrekende waarden in relatie tot afnameregio voor het jaar 2023. Voor een volledig overzicht van alle regio's met het hoogste percentage missings op het desbetreffende item, ook wat betreft het jaar 2022, zie bijlage 2.

Tabel 3.8 Top 3 hoogste percentages ontbrekende waarden per item gezien regio van afname

Domein - Item	% missings	Afnameregio	Statistiek (<i>p</i> -waarde, χ^2)
Relaties — Verskil contacten online offline	64,6%	Haarlem	$p<0,001$, $\chi^2=384,36$
Gezin — Opleidingsniveau vader	57,5%	Amsterdam	$p<0,001$, $\chi^2=184,34$
Werk, Vrije Tijd, Financiën — Online tijdsbesteding	48,3%	Haarlem	$p<0,001$, $\chi^2=386,67$

Instrumenttype

Tot slot is ook onderzocht of er significante verbanden bestaan tussen het aantal ontbrekende waarden en het instrumenttype, namelijk of het een reguliere Ritax A of Ritax B-afname, versus een verkorte IVS-2A-afname betreft. Voor beide jaren werd onder meer dan de helft van de items een statistisch significant verband gevonden tussen de mate van ontbrekende waarden en het instrumenttype. Er was echter geen eenduidige richting in dat verband. Soms ontbrak informatie juist significant vaker bij de verkorte, IVS-2A-afname, en soms juist significant vaker bij de reguliere Ritax A- of Ritax B-afnames. In de interviews en focusgroepen werd opgemerkt dat verkorte IVS-afnames waarschijnlijk minder volledig (kunnen) worden ingevuld, onder andere door praktische complicaties bij dit type afname. Zoals, tijdsdruk en personele beperkingen, vooral wanneer afnames plaatsvinden in weekenden of buiten reguliere werkdagen. Voor een volledig overzicht van alle items met een significante verband en de richting daarvan voor beide jaren, zie bijlage 2.

3.3.5

Tussenconclusie

De kwantitatieve analyse van ontbrekende waarden op itemniveau laat zien dat het percentage daadwerkelijke, onverklaarbare lege cellen beperkt is. Onverklaarbare lege cellen zijn lege velden die wel ingevuld hadden moeten zijn, terwijl 'onbekend'-scores ingevulde velden zijn waarin letterlijk 'onbekend' stond genoteerd. Hier is dus iets ingevuld, maar het daadwerkelijke antwoord op de vraag is niet bekend. Uit de ontbrekende waarden analyses bleken relatief hoge percentages 'onbekend'-scores bij bepaalde items. Als een strikte drempelwaarde van maximaal 5% ontbrekende waarden per item gehanteerd wordt, valt bij de meeste domeinen (ruim) meer dan de helft van de items af. Met name binnen de domeinen gezin, attitude en relaties gaat het om relatief hoge percentages van ontbrekende waarden. Als de meest ruime drempelwaarde van maximaal 20% ontbrekende waarden wordt gehanteerd, blijft voor alle domeinen ongeveer drie kwart van de items behouden. Met uitzondering van het domein relaties, waar dan alsnog 33% van de items afvalt.

De resultaten van 2022 in vergelijking met 2023 lijken te suggereren dat het aantal ontbrekende waarden over de tijd iets is verminderd. Echter, voor trendanalyses over de tijd zal rekening moeten worden gehouden met de structurele mate van ontbrekende waarden binnen sommige items en domeinen. Verder blijkt uit de subgroep analyses dat de ontbrekende waarden niet willekeurig verdeeld zijn. Binnen het domein gezin, school en attitude komen ontbrekende waarden vooral voor bij jongens, terwijl meisjes vaker ontbrekende waarden hebben op sommige items binnen het domein geestelijke gezondheid. Ontbrekende waarden komen ook structureel vaker voor bij oudere jongeren, vooral op schoolgerelateerde items. Verder zijn er duidelijke regionale verschillen. Bepaalde regio's (waaronder Amsterdam en Haarlem) scoren consequent hoger op ontbrekende waarden, wat suggereert dat registratiepraktijken en werkwijzen regionaal kunnen verschillen.

Samengevat bevestigen de kwantitatieve analyses de resultaten uit de interviews en focusgroepen (zie ook paragraaf 3.2). Naast inhoudelijke variatie in de werkwijze van raadsonderzoekers in de praktijk zijn er dus ook structurele verschillen in het voorkomen van ontbrekende waarden. Voor bepaalde domeinen en items, en binnen specifieke subgroepen, wordt dit voor beide jaren signaleerd. Dit kan consequenties hebben voor de bruikbaarheid van de data voor onderzoeksdoeleinden, met name wat betreft analyses over de tijd.

Tabel 3.9 Tussenconclusie

Zijn er indicatoren in de datasets welke duiden op:	Bevinding
Problematische ontbrekende waarden in specifieke domeinen of items?	Ja - Verschillende items met meer dan 5%, 10%, of 20% ontbrekende waarden - Soms vallen veel items binnen bepaalde domeinen af
Structurele ontbrekende waarden binnen bepaalde subgroepen?	Ja Leeftijd, geslacht en regio van afname lijken soms een rol te spelen in de mate van ontbrekende waarden

Zoals eerder benoemd in paragraaf 3.1.4, blijven er op basis van de kwalitatieve resultaten in deze haalbaarheidsstudie, 50 items over die over de tijd ongewijzigd zijn gebleven over de tijd (zie ook tabel 3.1). Als in het kader van betrouwbaarheid deze 50 items getoetst worden aan de meest strikte drempel van maximaal 5% ontbrekende waarden, blijven nog 18 (van de in totaal 108) items over, zoals weergegeven in tabel 3.10. Dit zijn dus de items die overblijven, als men voor onderzoeksdoeleinden alléén gebruik zou willen maken van ongewijzigde items, en tevens voldoende gevulde items over de tijd.

Tabel 3.10 Ongewijzigde en voldoende ingevulde* items over de tijd

Domein en Itemnummer	Itemnaam
Gezin	
1.2	Pleeggezin- en/of tehuisverleden
1.4a	Ouder(s)/verzorger(s) stelt/stellen regels op
1.4b	Ouder(s)/verzorger(s) ziet/zien toe of de jongere zich aan de regels houdt
1.5	Gehoorzaamheid van de jongere aan ouders/verzorgers (de mate waarin de jongere zich aan afspraken en regels houdt)
1.9	Is de jongere wel eens door een GI (Gecertificeerde Instelling) voor JB en JR vrijwillig of een kinderrechter civielrechtelijk uit huis geplaatst?
School	
2.1	Aantal dagen dagbesteding (school, stage en/of werk) per week.
2.4a	Schoolsituatie jongere afgelopen 6 maanden
2.5	De jongere hecht belang aan onderwijs/opleiding
2.9b	Kruis aan welke actie heeft plaatsgevonden [bij gedragsproblemen]
2.12b	Leeftijd bij de eerste schorsing in zijn/haar leven
Werk, vrije tijd, financiën	
3.5	Huidige interesse in een baan
Geestelijke Gezondheid	
6.2a	Is er een diagnose van psychische problemen die nu nog geldig is?
6.2c	Is medicatie voorgeschreven?
6.3	Staat de jongere onder behandeling van een psycholoog/psychiater?
6.4	Stemming
Agressie	
8.1	Mate waarin jongere vindt dat schreeuwen en verbale agressie geschikt zijn voor het oplossen van een conflict
8.2	Mate waarin jongere vindt dat vechten en fysieke agressie geschikt zijn voor het oplossen van een conflict
Vaardigheden	
9.2	Impulsiviteit, handelen alvorens na te denken

* Gezien de strikte drempel van minder dan 5% ontbrekende waarden

3.4 Interviews en focusgroepen — Data governance-vraag

Zoals beschreven in paragraaf 2.2.2 hebben twaalf professionals (van het ministerie van JenV en van het WODC) deelgenomen aan de focusgroepen ten behoeve van een data governance-vraag. De eerste, interne focusgroep met alleen WODC-medewerkers richtte zich met name op de noodzaak en behoefte, grondslag, privacyrisico's en de (benodigde) infrastructuur in het geval van een structurele levering van gegevens op microniveau uit het LIJ aan het WODC. De externe focusgroep concentreerde zich op technische en praktische mogelijkheden versus beperkingen omtrent een structurele levering of afslag van de data, de inrichting van bijbehorend databeheer, en de benodigde afstemming en samenwerkingsverbanden tussen relevante ketenpartners. De resultaten zijn hieronder samengevoegd onder het onderwerp inhoudelijke noodzaak, hoewel enkele andere aspecten ook ter sprake komen. Vooraf is ook opgemerkt dat alle opties en (on)mogelijkheden die worden beschreven een preliminaire inventarisatie betreffen die naar alle waarschijnlijkheid incompleet is.

3.4.1 Inhoudelijke noodzaak

Sommige kennisvragen op het terrein van jeugdcriminaliteit en de aanpak daarvan kunnen baat hebben bij de beschikbaarheid van LIJ-data omdat het een rijke, aanvullende databron kan zijn voor interne onderzoeken. Bijvoorbeeld, als aanvulling op het monitoren van ontwikkelingen in jeugdcriminaliteit, om de ontwikkelingen van risicoprofielen over de tijd beter te duiden, of ter versteviging van (quasi-)experimentele onderzoeken naar de effecten van interventies. Echter, deze behoefte is vooral projectmatig en wisselend over de tijd. Dit geldt zowel voor onderzoek dat intern wordt uitgevoerd als extern wordt uitgezet. Op dit moment ontbreekt daardoor de inhoudelijke noodzaak voor het structureel binnenhalen van LIJ-data op microniveau door het WODC. Een structurele levering van LIJ-data op microniveau vereist namelijk meer dan een inhoudelijke wens of projectmatige behoefte: het veronderstelt een continue en stabiele behoefte aan actuele en volledige gegevens op individueel niveau. Denk bijvoorbeeld aan een situatie waarin structurele beleidsmonitoring nodig is, waarbij periodiek inzichten gewenst zijn in kenmerken, trajecten of risicoprofielen van jongeren. Ook bij systematische effectevaluaties van landelijk beleid of doorlopende trendanalyses over meerdere jaren zou een structurele beschikbaarheid van gegevens uit het LIJ op microniveau mogelijk verdedigbaar zijn. Zolang er echter geen sprake is van een dergelijke continue informatiebehoefte, volstaat het om de data op ad-hocbasis op te vragen binnen de kaders van een concreet onderzoek.

Wat betreft verdere privacy- en AVG-vraagstukken zal in de toekomst ook rekening moeten worden gehouden met een aantal andere technische en praktische overwegingen. Als de data aanvraag LIJ-data uit de gehele jeugdstrafrechtketen betreft, zal formele toestemming van alle LIJ-data eigenaren verkregen moeten worden. Wat de Ritax-data betreft, gaat dit specifiek om de RvdK én alle individuele, gecertificeerde instellingen. Deze toestemmingsprocedures nemen tijd in beslag. Aan de andere kant zou er ook voor gekozen kunnen worden om alleen Ritax-data van de RvdK te werven, waardoor geen toestemming nodig is van andere formele eigenaren. Daarnaast vereist een structurele levering van LIJ-data ook afstemming met de privacyofficers van alle betrokken partijen. Hierin moeten specifieke privacy en AVG-vragen dan verder moeten worden uitgewerkt met oog op het feit dat het gaat om gegevens van een kwetsbare onderzoeksdoelgroep (minderjarigen en strafrechtelijke RvdK).

Daarnaast is het ook belangrijk de beperkingen van de LIJ-data voor onderzoeksdoeleinden te overwegen. De LIJ-gegevens zijn primair bedoeld voor werkzaamheden van justitiële ketenpartners en niet ontwikkeld voor onderzoek. Sommige respondenten benoemden dat de manier waarop LIJ-data geregistreerd wordt (logischerwijs) niet hoofdzakelijk is ingericht voor onderzoeksdoeleinden, en dat LIJ-data veel ontbrekende waarden bevatten waardoor de dekking per domein of item beperkt zou zijn. Dergelijke zorgen omtrent de bruikbaarheid van LIJ-data voor onderzoeksdoeleinden werden ook genoemd in de externe focusgroep.

Deelnemers aan de externe focusgroep benoemden dat uit de structurele levering van LIJ-data tussen Justid en DGSenB meerdere beleidsproducten zijn voortgekomen (voor zover bekend niet openbaar toegankelijk). Volgens enkele deelnemers kon dit toegeschreven worden aan het feit dat de datalevering zorgvuldig afgestemd is op het specifieke onderzoeksdoeleinde van DGSenB. Namelijk, het monitoren van LIJ-afnames en het beantwoorden van daaruit voortkomende beleidsvragen. Het lijkt dus mogelijk om voor beleidsdoeleinden een structurele levering van LIJ-data te verkrijgen. Echter het is de vraag of dit ook kan voor onafhankelijk wetenschappelijk onderzoek en of structurele toegang tot LIJ-data daarvoor wel noodzakelijk is. In toekomstige overwegingen voor het gebruik van LIJ-data voor onderzoeksdoeleinden moet ook onderscheid gemaakt worden tussen incidenteel, projectmatig gebruik van LIJ-data versus structureel beheer van LIJ-data, en de herleidbaarheid van de benodigde gegevens. Structureel beheer brengt grote verantwoordelijkheden en risico's met zich mee: de verantwoordelijk onderzoeker(s) moeten weten wat de data betekenen, hoe deze verandert over de tijd, en hoe deze data beveiligd moeten worden opgeslagen. Als de LIJ-data gekoppeld moeten worden, dan geldt een verhoogd privacy risico wat stevige beveiligingsmaatregelen vraagt.

Structurele levering van LIJ-data is technisch realiseerbaar. Dit blijkt bijvoorbeeld al uit het bestaande samenwerkingsverband tussen Justid en DGSenB. Echter, een dergelijke dataleveringsovereenkomst vraagt naast een duidelijke behoefteformulering ook goed ingerichte processen. Justid levert momenteel maandelijks LIJ-data aan Ketenregie en geeft aan dat een jaarlijkse of vijfjaarlijkse levering richting het WODC inpasbaar is. Echter, dan moet onder meer helder zijn voor welke onderzoeksdoeleinden de LIJ-data benodigd is en op welk niveau de data wenselijk is (bijv. geaggregeerde of microgegevens).

Een structurele levering vraagt ook om organisatorische afspraken over wie 'wat' doet. Op dit moment is het beheer van LIJ-data bij Justid belegd. Verschillende respondenten benoemden dat de huidige LIJ-database gedateerd en weinig gestandaardiseerd is en soms zelfs handmatige databewerkingen vereist. Derhalve is vanuit Justid aangestuurd op 'ruwe' dataleveringen in toekomstige dataleveringsovereenkomsten, met zo weinig mogelijk tussenschakels. Dit zou betekenen dat er meer verantwoordelijkheden belegd moeten worden bij de ontvangende partij. Die partij zou naast beheer dan ook zelf moeten bijhouden hoe de LIJ-data geïnterpreteerd moeten worden en welke fouten of onvolledigheden in een databestand zitten. Dit lijkt in te gaan tegen de huidige richting die vanuit JenV is ingezet in het algemene GegevensLeveringProtocol (GLP) waarbij broneigenaren meer verantwoordelijk worden gesteld voor de kwaliteit van de data.

Binnen het WODC is afgesproken dat de afdeling Informatievoorziening & Datamanagement (I&D) het formele aanspreekpunt is voor databeheer, dataverzoeken en datastewardship. De inhoudelijke duiding ligt bij de kennislijnen. In het geval van

een structurele levering van LIJ-data (met beheer intern belegd bij het WODC) is een samenwerking nodig tussen I&D, en onderzoekers, waarbij I&D de datastromen zou beheren, en onderzoekers de inhoudelijke duiding leveren. Dit vergt veel capaciteit. De exacte capaciteit die een dergelijke structurele levering vraagt, is op dit moment niet in te schatten. In het geval van weinig veranderingen aan de LIJ-data over de tijd is beheer relatief beperkt. Dit scenario zal zich naar verwachting niet snel voordoen, uit het voorgaande bleek al dat er vaak wijzigingen plaatsvinden. Bij fouten of wijzigingen en aanpassingen aan de LIJ-data moet er snel kunnen worden geschakeld en aangepast. Daarnaast zijn de LIJ-data bedoeld als ondersteunend voor de praktijk en niet in eerste instantie ingericht voor onderzoeksdoeleinden. Er zijn dus structurele controles nodig op de kwaliteit van opgevraagde LIJ-data. Dit kost veel tijd, zoals bij het structureel beheer van andere databronnen binnen het WODC reeds is gebleken en ook blijkt uit de inspanningen verricht in dit haalbaarheidsonderzoek. Uit de interne focusgroep bleek dat het WODC de nodige infrastructuur heeft voor het beheren van gevoelige data (zoals bij OBJD of RAC-min data), maar LIJ-data is vooralsnog onbekend terrein en zal uitbreiding van deze infrastructuur vragen. Er zal ook expertise moeten worden opgebouwd om de LIJ-data te leren begrijpen, de kwaliteit (periodiek) te controleren en deze betrouwbaar te koppelen aan andere bronnen.

Tot slot is er een financiële vraag. Structureel beheer van dergelijke data kost personele inzet, en uitbreiding en inzet van (ICT)systemen. Indien in de toekomst ook overdracht van LIJ-data aan derden zou worden overwogen, kan het noodzakelijk zijn om vooraf stil te staan bij een kostendekkend model.

3.4.2 *Tussenconclusie*

Samenvattend wijst de informatie hierboven erop dat een inhoudelijk noodzaak voor het structureel leveren van LIJ-data op microniveau aan het WODC vooralsnog ontbreekt. Dit gebrek aan noodzaak betekent een onvoldoende juridische grondslag voor de structurele levering van microdata aan en beheer ervan door het WODC. Ook is het de vraag hoeveel beheercapaciteit een structurele levering vergt voor het WODC. Technisch en organisatorisch lijken er wel mogelijkheden, maar dit vraagt de nodige opbouw in capaciteit en expertise. Daarbij is een belangrijke vraag in hoeverre het databaseer bij de dataeigenaar of dataleverancier (Justid) blijvend kan worden belegd en hoeveel taken tot de verantwoordelijkheid van de data aanvrager (het WODC) zullen behoren. Daarbij speelt mee dat de kwaliteit van de LIJ-data vooralsnog niet structureel consistent op orde lijkt te zijn, een verantwoordelijkheid die vanuit de huidige insteek van de GLP bij de broneigenaren ligt. Totdat deze zaken omtrent noodzaak, databaseer en datakwaliteit meer duidelijkheid verkrijgen, ligt de meest realistische optie in een goed afgestemde, projectmatige levering van LIJ-data.

4 Slot

4.1 Beantwoording onderzoeksvragen

In deze haalbaarheidsstudie stond de vraag centraal in welke mate gegevens verzameld in het Landelijke Instrumentarium Jeugd (LIJ) en meer specifiek de Ritax-gegevens geschikt zijn wetenschappelijk onderzoek om (structureel) ontwikkelingen in kenmerken (onder meer risico- en beschermende factoren) van jeugdige verdachten over de tijd te kunnen vaststellen. Daartoe zijn in dit rapport verschillende deelvragen onderzocht, die hieronder beknopt worden beantwoord.

Onderzoeksvraag 1a betrof de mate waarin gegevens van het LIJ door de tijd heen op een consistente wijze zijn geregistreerd. Oftewel, de mate waarin de informatie vanuit een raadsonderzoek op vergelijkbare wijze in het IFM-systeem wordt geregistreerd. Uit de literatuurstudie, interviews en focusgroepen blijkt dat de consistentie in registratie van gegevens uit het LIJ over de tijd wordt beperkt door meerdere factoren. In de eerste plaats leiden inhoudelijke wijzigingen die in de afgelopen jaren aan (delen van) het instrument zelf zijn gedaan, met name de introductie van de Ritax 2.0 in 2022, tot een methodebreuk die de vergelijkbaarheid vóór en na dit moment bemoeilijkt. Daarnaast bestaan in de praktijk verschillen in de wijze van afname en registratie, zowel tussen regio's als tussen raadsonderzoekers. Dit beeld wordt bevestigd door de kwantitatieve analyses, die structurele verschillen in ontbrekende waarden laten zien tussen domeinen, regio's en subgroepen. Verder blijkt uit de kwantitatieve vergelijking van Ritax gegevens tussen twee jaren dat sommige items structureel hogere percentages ontbrekende waarden hebben in het ene jaar, ten opzichte van het andere jaar. Samengevat kan worden geconcludeerd dat de data niet volledig consistent geregistreerd zijn over de tijd. Dit beperkt de wetenschappelijke bruikbaarheid van de gegevens voor vergelijkingen over de tijd. Dit geldt met name voor data van vóór de Ritax 2.0-periode, gezien de significante aanpassingen die daarmee zijn doorgevoerd.

Onderzoeksvraag 1b richtte zich op de aard en de achtergrond van eventuele veranderingen in de wijze van registratie. De resultaten uit het kwalitatieve gedeelte laten zien dat veranderingen in de wijze van registratie voornamelijk voortkomen uit inhoudelijke aanpassingen aan het instrument zelf en variatie in de wijze van afname in de praktijk. De overgang naar de Ritax 2.0 heeft geleid tot een verandering in de opbouw van domeinen en items, waardoor scores en totaalscores op een andere manier tot stand komen dan in eerdere versies. Daarnaast blijkt uit de interviews en focusgroepen dat raadsonderzoekers verschillen in ervaring en interpretatie meebrengen, wat invloed kan hebben op de wijze van registratie. Veranderingen in de registraties zijn dus niet alleen inhoudelijk methodisch, maar hangen ook samen met de uitvoeringspraktijk.

Onderzoeksvraag 2a ging in op de mate waarin gegevens van de Ritax voldoende zijn ingevuld over de tijd. De kwantitatieve analyses van ontbrekende waarden laten zien dat de data niet altijd volledig zijn ingevuld. Hoewel het aantal daadwerkelijk lege cellen beperkt is, wordt op bepaalde items regelmatig 'onbekend' gescoord. Dergelijke 'onbekend' scores waren het meest prominent aanwezig binnen de domeinen gezin, attitude en relaties. Andere domeinen hadden soms ook veel items met 'onbekend'-scores, afhankelijk van of een cut-offwaarde van maximaal 5%, 10% of 20% ontbrekende waarden werd gehanteerd. Een vergelijking tussen de jaren 2022 en

2023 suggereert dat de mate van ontbrekende waarden enigszins is afgenomen over de tijd, maar dat ontbrekende waarden blijven bestaan. Dit betekent dat volledigheid van registratie van de data niet vanzelfsprekend is. Dit betekent dat de bruikbaarheid van de data voor onderzoeksdoeleinden, zoals analyses over de tijd, in zekere mate beperkt is, en dat dit mede afhangt van de wijze waarop met de ontbrekende waarden wordt omgegaan.

Onderzoeksvraag 2b bouwt hierop voort en betrof de verschillen tussen domeinen en items in de mate van ontbrekende waarden. De kwantitatieve analyses wijzen erop dat ontbrekende waarden niet gelijkmatig verdeeld zijn over domeinen en items. Zoals hierboven benoemd, hebben de domeinen gezin, attitude en relaties relatief hoge percentages van ontbrekende waarden, terwijl andere domeinen vollediger zijn ingevuld. Ook blijkt dat ontbrekende waarden niet willekeurig voor komen. Bij jongens wordt vaker 'onbekend' gescoord op gezins- en school gerelateerde items, en bij meisjes juist op sommige items binnen het domein geestelijke gezondheid. Verder blijkt dat voor oudere jongeren en in bepaalde regio's, waaronder Amsterdam en Haarlem, structureel hogere percentages van ontbrekende waarden zijn. Deze bevindingen suggereren dat verschillen in registratie tussen domeinen, subgroepen en regio's mogelijk systematisch van aard zijn. Echter, op basis van gegevens over twee jaren kan dit niet met zekerheid worden vastgesteld. Dit heeft wel gevolgen voor de bruikbaarheid van data voor vergelijkingen over de tijd omdat niet duidelijk wordt in hoeverre de jaardata al dan niet selectief zijn en daarmee de representativiteit beperken.

Onderzoeksvraag 3 ging over het moment vanaf wanneer de gegevens geschikt zijn om te gebruiken voor onderzoek naar ontwikkelingen in kenmerken van de populaties over de tijd. Op basis van de bevindingen in dit rapport kan worden geconcludeerd dat de data voor dergelijke vergelijkingen bruikbaar zijn vanaf de invoering van Ritax 2.0 in 2022 dan daarvoor. We maken hierbij wel direct een kanttekening. Ook na de invoering van Ritax 2.0 blijven er aandachtspunten over de registratie, zoals de hoge mate van 'onbekend'-scores op sommige items en de aanwezigheid van structurele patronen in ontbrekende waarden. Zoals hierboven benoemd, is het voor analyses met recentere gegevens dus belangrijk rekening te houden met de aanwezigheid van (structurele) ontbrekende waarden en de mogelijke gevolgen daarvan voor de conclusies in een dergelijk onderzoek. Het is op basis van de geleverde data namelijk niet duidelijk in hoeverre hiermee op jaarbasis een selectief beeld van de onderzoeksgroepen ontstaat. Koppeling aan een andere bron waarin de volledige populatie van jeugdige verdachten of justitiabelen aanwezig zijn, zou hier duidelijkheid in kunnen bieden.

Onderzoeksvraag 4 richtte zich op veranderingen in de populatie van jeugdige verdachten waarbij een Ritax-afname is over de tijd, met betrekking tot hun demografische en delict kenmerken. Vanwege de methodeverandering naar de Ritax 2.0 is deze vraag slechts voor twee opeenvolgende jaren beantwoord. De analyses van de beschrijvende statistieken voor 2022 en 2023 laten zien dat de leeftijd van de jongeren in beide jaren vrijwel gelijk is, met een gemiddelde leeftijd van 15,4 jaar. In beide jaren zijn jongens in de meerderheid met ongeveer 90% jongens en 10% meisjes. De regionale spreiding van deze LIJ-afnames is ook ongeveer gelijk voor beide jaren, met Rotterdam, Den Haag, en Arnhem als de drie regio's met meeste afnames. Bij delict zwaarte is wel een lichte verschuiving gevonden, maar er blijven in beide jaren te veel onbekende waarden zodat we hierover verder niets kunnen zeggen.

Ondanks de kleine verschillen lijkt de populatie tussen 2022 en 2023 nauwelijks veranderd wat betreft deze kenmerken.

Onderzoeksvraag 5 betrof mogelijke aanpassingen die noodzakelijk of haalbaar zijn om de kwaliteit van geregistreerde LIJ-gegevens te verbeteren voor onderzoek naar vergelijkingen over de tijd. De bevindingen bieden verschillende aanknopingspunten. Allereerst kan meer uniformiteit in de afname en de registratie bijdragen aan de betrouwbaarheid van gegevens uit het LIJ. Daarnaast kan het terugdringen van 'onbekend'-scores, met name binnen domeinen waarin dit veel voorkomt (gezin, attitude/houding en relaties), de bruikbaarheid van de data voor vergelijkingen over de tijd vergroten. Ook is het belangrijk dat inhoudelijke wijzigingen aan het instrument zorgvuldig en structureel worden gedocumenteerd, zodat deze kunnen worden getraceerd en veranderingen kunnen worden geduid.

Zoals benoemd in paragraaf 1.1 konden de onderzoeksvragen 6 t/m 9, welke betrekking hebben op een overkoepelende data governance vraag, in dit onderzoek ten dele worden beantwoord. De beantwoording van deze onderzoeksvragen is namelijk afhankelijk van een positieve beantwoording van de betrouwbaarheidsvraag. Desondanks zijn tijdens twee focusgroepen ten behoeve van dit onderwerp wel mogelijkheden en beperkingen gesignaleerd. Daaruit kwam naar voren dat voor het structureel binnenhalen en verwerken van gegevens uit het LIJ op microniveau niet alleen technische en organisatorische afstemming vereist is, maar ook een duidelijke en blijvende informatiebehoefte. Een structurele levering zou pas aan de orde zijn wanneer er sprake is van een continue behoefte aan actuele gegevens. Bijvoorbeeld, voor het monitoren van ontwikkelingen in populaties ten behoeve van beleid, vergelijkbaar met bijvoorbeeld het monitoren van jeugdcriminaliteit of recidive als relevante indicatoren van jeugdbeleid. Momenteel ontbreekt een dergelijke inhoudelijke noodzaak voor structurele levering van gegevens uit het LIJ op microniveau aan het WODC. Daarbij komt dat uit de analyses blijkt dat de mate waarin de data zijn geregistreerd beperkingen kent en niet consistent is over verschillende regio's. Tot slot is genoemd dat er wel technische en organisatorische mogelijkheden zijn voor levering van microdata, maar dat structurele leveringen van LIJ-data (personele en materiële) capaciteit vergt die niet zonder meer beschikbaar is.

Samenvattend blijkt uit deze haalbaarheidsstudie dat het op dit moment *niet* voldoende verantwoord is om LIJ-data *structureel* te gebruiken voor wetenschappelijk trendonderzoek. Vooralsnog vertonen deze gegevens te veel variatie in registratie en invulling om betrouwbare vergelijkingen te maken over de tijd. Mits de datakwaliteit wordt versterkt, kan structureel gebruik op termijn mogelijk wél verantwoord zijn. Voor *incidenteel* onderzoek naar vergelijkingen over meerdere jaren kunnen de gegevens uit het LIJ mogelijk *wel* bruikbaar zijn waarbij het belangrijk is dat ze gekoppeld worden aan een externe bron die gegevens van alle jeugdige verdachten of justitiabelen omvat. Daarbij moet dan ook zorgvuldig worden omgegaan met – en gerapporteerd over – de mate van ontbrekende waarden en de gevolgen daarvan voor de conclusies.

4.2 Mogelijkheden en beperkingen van het onderzoek

Deze haalbaarheidsstudie kent enkele beperkingen en biedt tegelijkertijd nieuwe inzichten over het gebruik van de Ritax voor onderzoek naar vergelijkingen van de kenmerken van de populatie jeugdige verdachten over de tijd. Allereerst is een

belangrijke beperking dat het onderzoek geen volledig inzicht heeft kunnen bieden in de achterliggende redenen *waarom* in de praktijk sommige items structureel minder worden ingevuld. Uit de kwalitatieve bevindingen blijkt wél dat verschillen in registratie onder meer samenhangen met variatie in afname en interpretatie van raadsonderzoekers. Daarnaast lijkt het voor sommige items in de praktijk moeilijk om betrouwbare informatie te verkrijgen, bijvoorbeeld bij vragen over het sociale netwerk van jongeren. Bij het domein 'relaties' kan het bijvoorbeeld lastig zijn om informatie te achterhalen wanneer jongeren zich beroepen op hun zwijgrecht of terughoudend zijn over het delen van informatie over antisociale vrienden. Het blijft echter onduidelijk waarom sommige items vaker op 'onbekend' worden gescoord dan andere.

Een tweede beperking is dat de empirische kwantitatieve analyses in dit onderzoek zich beperkten tot twee jaren (2022 en 2023). Daardoor is slechts een momentopname verkregen van de consistentie en volledigheid van de data. Of de waargenomen patronen in ontbrekende waarden tussen de items, domeinen en subgroepen structureel zijn, kon binnen dit onderzoek niet worden vastgesteld. Ook kon niet worden vastgesteld of deze patronen zich over een langere periode voordoen. In onderzoek over meerdere jaren zou het wellicht nuttig zijn om te bepalen of de resultaten van dit onderzoek terug te zien zijn of incidenteel waren.

Een derde punt is dat selectiviteit van de onderzochte populaties niet duidelijk is geworden. De kwantitatieve analyses geven wel inzicht in verschillen in ontbrekende waarden naar geslacht, leeftijd en regio. Echter, ze geven geen inzicht in de representativiteit van deze jongeren voor de algehele groep jeugdige verdachten of justitiabelen. Voor dergelijke analyses is koppeling met andere databronnen, zoals de politieregistratie BVH (beschikbaar bij het CBS) of de OBJD (bij het WODC) noodzakelijk. Zonder een dergelijke koppeling blijft onduidelijk in hoeverre de onderzochte groep een sub selectie vormt en welke implicaties dat heeft voor de generaliseerbaarheid van de bevindingen over hun kenmerken naar de totale populatie van jeugdige verdachten of justitiabelen.

Desondanks biedt deze haalbaarheidsstudie aanknopingspunten voor de bruikbaarheid van gegevens uit het LIJ en meer specifiek de Ritax voor onderzoek. Ten eerste biedt dit onderzoek inzicht in de (in)consistentie van invulling van afnames op itemniveau, voor meerdere recente jaren. Deze inzichten bieden ook aanknopingspunten voor de praktijk, bijvoorbeeld om te begrijpen waar knelpunten ontstaan bij de afname en registratie. Daarnaast maakt dit onderzoek inzichtelijk dat de vulling van de gegevens selectief kan zijn. Eerdere beleidsanalyses op basis van de data op itemniveau schetsen dus mogelijk een selectief beeld van de doelgroep van jeugdige justitiabelen. Het is echter niet duidelijk welke subgroepen precies over- of ondervertegenwoordigd zijn.

4.3 Discussie

De resultaten uit deze haalbaarheidsstudie bieden inzicht in de bruikbaarheid van gegevens uit het LIJ, specifiek de Ritax, voor vergelijkingen van kenmerken van jeugdige verdachten over de tijd. Daarnaast laten ze zien welke aandachtspunten belangrijk zijn bij het gebruik van de data voor beleid en onderzoek. Op basis van deze inzichten kunnen tevens aanbevelingen worden gedaan.

Een eerste thema betreft de inhoudelijke verandering van het instrument wat een methodebreuk met zich meebrengt. De invoering van de Ritax 2.0 in 2022 bracht significante inhoudelijke veranderingen in items, domeinen en totaalscores. Hierdoor zijn de gegevens van voor deze periode niet een-op-een vergelijkbaar met die van daarna. Dit beperkt het uitvoeren van betrouwbare vergelijkingen over de tijd. Het betekent dat oudere data voor 2022 voorzichtig moeten worden geïnterpreteerd wanneer ze worden vergeleken met gegevens van na die periode. Dat is niet zonder meer mogelijk. Het impliceert dat het gebruik van LIJ-data, in het bijzonder de Ritax, als aanvullende bron naar ontwikkelingen in kenmerken van jeugdige verdachten over de tijd pas goed mogelijk is vanaf 2022. Daarbij is het wel belangrijk dat het nodige wordt gedaan aan het verbeteren van de registratie van de gegevens (zie hierna). Om voor de toekomst de vergelijkbaarheid tussen jaren te verbeteren, is het belangrijk dat veranderingen in het instrument zorgvuldig en systematisch worden gedocumenteerd en dat nieuwe items of domeinen expliciet worden gelinkt aan eerdere metingen. Voor een betrouwbare vergelijking over de tijd is meer nodig. Belangrijk is om per jaar zicht te krijgen op de selectiviteit van de populatie jongeren waarbij een Ritax is afgenomen (bijvoorbeeld in hoeverre bepaalde groepen jongeren wel of niet in de geregistreerde gegevens voorkomen). Een betrouwbare vergelijking van jaren is bijvoorbeeld mogelijk door voor meerdere jaren vanuit een complete bron van jeugdige verdachten als de BVH of van justitiabelen als de OBJD een representatieve steekproef te trekken en op microniveau de items van de Ritax te gebruiken die over de tijd niet zijn veranderd.

Een tweede thema betreft de volledigheid van de gegevens van de Ritax en de aanwezigheid van ontbrekende waarden. Vooral in de domeinen gezin, attitude en relaties zijn (relatief) veel items met 'onbekend'-scores, waardoor sommige domeinen bij een strenge cut-off van maximaal 5% ontbrekende waarden bijna volledig onbruikbaar zijn voor adequate vergelijkingen. Daarbij blijkt dat ontbrekende waarden op sommige items systematisch vaker voorkomen bij specifieke subgroepen (uitgesplitst naar geslacht, leeftijd of regio van afname). Dit betekent dat het kritisch beoordelen van de data en zorgvuldig omgaan met ontbrekende waarden essentieel is, niet alleen door onderzoekers, maar zeker ook voor beleidsmedewerkers die de gegevens gebruiken voor beleidsramingen. Door vooraf te bepalen welke items, domeinen, of bijvoorbeeld regio's, worden meegenomen en door strategieën voor ontbrekende waarden vast te stellen (zoals het gebruik van alleen volledig ingevulde items, imputatiemethoden of sensitiviteitsanalyses) kan een vergelijking enigszins verbeteren, hoewel dit de mogelijke selectiviteit van de ingevulde gegevens niet oplost. Ook kan de verdeling van ontbrekende waarden per subgroep bekeken worden om te beoordelen of bepaalde groepen systematisch onder- of oververtegenwoordigd zijn in de datasets.

In het onderzoek zijn meerdere drempelwaarden onderzocht voor de mate waarin items ontbrekende waarden lieten zien. Het onderscheid tussen de drempelwaardes van maximaal 5%, 10% en 20% ontbrekende waarden laat zien dat de bruikbaarheid van de data afhankelijk is van de gehanteerde strengheid. Naarmate een ruimere marge wordt toegestaan, vallen steeds minder items uit. Hoewel een strenge eis van maximaal 5% ontbrekende waarden voor vergelijkend onderzoek aan te bevelen is, hangt de uiteindelijke marge af van de voorliggende onderzoeksvraag en van de mogelijkheden om selectiviteit te corrigeren. Bovendien is het van belang om te bepalen of, en op welke manier, rekening kan worden gehouden met de in dit onderzoek gesignaleerde beperkingen van gegevens uit het LIJ.

Voor praktijktoepassingen op individueel casusniveau is dit mogelijk minder problematisch, maar bij het gebruik van de data voor onderzoeksdoeleinden op groepsniveau ligt selectiviteit voor de hand. Een aanbeveling voor onderzoek waarin de gegevens worden gebruikt om uitspraken te doen over de populatie van jeugdige verdachten, is om dergelijke selectiviteit te onderzoeken. Wat betreft het verbeteren van de volledigheid van toekomstige registraties van LIJ-gegevens zijn denkbare implicaties voor de praktijk bijvoorbeeld het geven van extra training of ondersteuning aan raadsonderzoekers waar nodig, of mogelijk verbeterde landelijke afstemming over het invullen van 'onbekend'-scores.

Tegelijkertijd is het relevant om kort te benoemen waarom Ritax gegevens tóch gebruikt willen worden. In beleid en praktijk kan mogelijk de behoefte zijn om signalen over ontwikkelingen in de populatie van jeugdige verdachten, bijvoorbeeld om effecten van interventies of veranderingen in werkwijzen tijdig te kunnen volgen. Ook wanneer de datakwaliteit beperkingen kent, kunnen Ritax gegevens – mits zorgvuldig geïnterpreteerd – indicatieve inzichten bieden voor monitoring en trendonderzoek.

Dit raakt aan een derde thema, namelijk regionale verschillen in werkwijzen. De bevindingen uit deze haalbaarheidsstudie laten zien dat werkwijzen verschillen tussen regio's en tussen ketenpartners onderling. Het is begrijpelijk en noodzakelijk dat deze variatie er is, want het onderstreept het belang dat raadsonderzoekers discretionaire ruimte hebben om maatwerk te leveren in onderzoeken onder individuele jeugdige verdachten. Dit heeft echter wel implicaties voor de bruikbaarheid van de data voor onderzoek op groeps- of populatieniveau. Vergelijkingen tussen regio's op basis van deze gegevens zijn daardoor beperkt (vanwege grote (on)betrouwbaarheidsmarges) terwijl aan de andere kant analyses binnen eenzelfde regio mogelijk wel voldoende bruikbaar zouden kunnen zijn. Voor beleid betekent het dus dat vergelijkingen tussen regio's over de kenmerken van de populaties met de nodige voorzichtigheid moeten worden bekeken. Een interbeoordelaarsbetrouwbaarheidsonderzoek waarin regio's worden vergeleken zou zicht op deze verschillen kunnen bieden. Uitkomsten van een dergelijk onderzoek kunnen inzicht bieden in items of domeinen die gevoelig zijn voor interpretatieverschillen. Daarnaast kan het aanknopingspunten bieden voor het aanpassen van systemen, landelijke richtlijnen, of trainingen. Zo kan de kwaliteit en vergelijkbaarheid van LIJ-data over regio's en tussen ketenpartners mogelijk worden verbeterd.

Een vierde thema betreft structurele levering van gegevens uit het LIJ aan het WODC in relatie tot privacy-issues en data governance. Er zijn technische en organisatorische mogelijkheden om data structureel op microniveau te leveren. Vooralsnog ontbreekt het echter ten eerste aan een inhoudelijke noodzaak om dergelijk gevoelige gegevens te leveren, iets wat vanuit de AVG wel een vereiste is. Daarnaast vergt structurele levering van microdata (personele en materiële) capaciteit, afspraken over (beveiliging van) databeheer en aandacht voor privacy- en AVG-aspecten. De opties daarvoor moeten goed worden verkend met aandacht voor kansen en risico's. Als deze worden uitgewerkt zijn duidelijke afspraken nodig tussen de verschillende (bron)beheerders en het WODC over de diverse onderdelen van de data governance.

Een alternatief en minder invasief scenario om inzicht te krijgen in kenmerken van jeugdige verdachten en veranderingen daarin over de tijd is bijvoorbeeld door om de paar jaar een onderzoek aan te vragen naar de populatie jongeren waarbij de Ritax (of breder ook het LIJ) is afgenomen. Hierbij is het dan ook belangrijk om rekening te houden met bovengenoemde punten over inhoudelijke veranderingen over de tijd,

(regionale en temporele) variaties in afnames, volledigheid van de gegevens en selectiviteit daarvan.

Tegelijkertijd is het van belang te beseffen dat beide vormen (structureel of incidenteel) de nodige materiële en personele capaciteit vragen, die bij structurele levering aanzienlijker zijn dan bij meer incidenteel onderzoek. Voordat het zover komt, is het in ieder geval aan te raden om tussentijds de documentatie en codeboeken van de Ritax en andere LIJ-gerelateerde instrumenten bij te (blijven) houden die alle wijzigingen en redenen voor aanpassingen aan het instrument en LIJ-datasets vastleggen. Dit kan de bruikbaarheid van toekomstige LIJ-data aanzienlijk verbeteren.

Tot slot, de bevindingen uit deze haalbaarheidsstudie bieden aanknopingspunten voor verbeteringen en aanbevelingen voor de toekomst. Een meer uniforme afname en registratie en meer inzicht in de motieven achter verschillende werkwijzen omtrent de instrumenten van het LIJ en de invulling ervan kunnen de betrouwbaarheid en consistentie van de data verbeteren. Vanuit beleidsmatig perspectief is het daarbij relevant om meer inzicht te krijgen in hoe het instrument in de praktijk precies wordt afgenomen. Daarbij is het belangrijk om te begrijpen waar in dat proces eventuele knelpunten optreden, en hoe die de kwaliteit en vergelijkbaarheid van de registraties over de tijd beïnvloeden. Het verminderen van (regionale en temporele) variatie in registratie en afname, het terugdringen van 'onbekend'-scores in items en domeinen met veel ontbrekende waarden, en het vergroten van de kennis en training van raadsonderzoekers waarin ook het algemene belang van volledige registraties aan de orde komt, kan bijdragen aan een betere kwaliteit van de geregistreerde data in het LIJ-instrumentarium. Dit geldt niet alleen voor de Ritax waar dit onderzoek voornamelijk op ziet, maar waarschijnlijk ook voor de andere instrumenten gebruikt in het LIJ. Voor de praktijk betekent dit dat het instrument betrouwbaarder wordt en dat beleids- en onderzoeksdoelen beter ondersteund kunnen worden. Het biedt betrouwbaardere data, die ingezet kunnen worden voor vergelijkingen over de tijd, subgroep vergelijkingen en evaluaties van interventies. Dat is niet alleen voor onderzoekers relevant, maar ook voor individuele jongeren en voor beleid dat het LIJ gebruikt voor sturing op haar inzet en doelen.

Samengevat blijkt uit deze haalbaarheidsstudie dat, hoewel gegevens uit het LIJ, en dan specifiek de Ritax, waardevolle inzichten zouden kunnen bieden om ontwikkelingen over de tijd in kenmerken van jeugdige verdachten te beschrijven, er op dit moment beperkingen zijn om dat *structureel* te doen. Zo is er in het recente verleden een methodebreuk door de overgang naar een nieuwe versie van het LIJ, is er regionale variatie in afname en zijn er aanzienlijke aantallen onbekende waarden, waardoor onduidelijk blijft of de beschikbare gegevens per jaar representatief zijn voor de groepen jeugdige verdachten. Dat beperkt vergelijken over de tijd, maar maakt het niet onmogelijk. *Incidenteel* onderzoek naar veranderingen in achtergrondkenmerken van jeugdige verdachten of veroordeelden over de tijd waarbij Ritax-gegevens over meerdere jaren worden gebruikt, lijkt wel uitvoerbaar, mits aandacht wordt besteed aan de in dit onderzoek geschetste beperkingen. Er moet dan wel zorgvuldig worden afgewogen hoe in de onderzoeksopzet om te gaan met (de mate van) ontbrekende waarden en de gevolgen daarvan voor de conclusies van dergelijk onderzoek. Om de selectiviteit van de informatie te beperken en uitspraken over de populatie van jeugdige verdachten te kunnen doen, is een koppeling nodig aan een externe bron waarin gegevens over de volledige populatie van jeugdige verdachten of justitiabelen beschikbaar zijn (denk aan de politieregistratie bij het CBS of de OBJD bij het WODC).

Summary

Juvenile suspects in sharper focus?

A feasibility study to determine whether the Ritax (risk assessment) data obtained from the National Screening and Risk Assessment Instrument for Juvenile Offences are sufficiently reliable to allow comparisons over time.

This feasibility study investigates the extent to which Ritax data from the Dutch Child Protection Board, which are part of the National Screening and Risk Assessment Instrument in the Juvenile Justice System (Landelijk Instrumentarium Jeugstrafrechtketen), are suitable for investigating changes in the characteristics of juvenile suspects over time. For example, as a supplementary source of data for the Juvenile Crime Monitor (JCM), which examines trends in juvenile crime every two years. The national police records that are available provide little information about the characteristics of juvenile suspects, such as relevant risk and protective factors that affect the incidence of delinquent behaviour. Although self-reported data, as measured in the Self-Reported Juvenile Delinquency Monitor, do provide this information, they focus on general populations. The Ritax data could potentially fill this gap for all juvenile suspects because of the screening for risk and protective factors that may play a role in either preventing criminality or promoting reoffending, as the case may be. However, it is not known how consistently the Ritax data are recorded nationwide and over time (information that would be necessary if comparisons are to be made over time). Such information is essential if statements are to be made about changes in the characteristics of juvenile suspects over time (or about differences between groups). This study has therefore focused on the extent to which the Ritax data can be used for comparisons over time of juvenile suspects at group and population level. This concerns the reliability of the Ritax data in terms of the manner in which these data were entered in the recording systems available nationwide and the completeness of the data.

A mixed-methods study was conducted to gauge the reliability and usability of the Ritax data. First, a literature study was carried out to determine what was already known from previous studies about the design, reliability and usability of the National Screening and Risk Assessment Instrument for Juvenile Offences (LIJ) in general and the Ritax data in particular. Next, the method of collecting and entering Ritax data and determining their usability in practice was examined through qualitative expert interviews and focus groups with child welfare investigators, policy officers and professional practitioners. Third, quantitative analyses of anonymised Ritax datasets were conducted to assess the entry and completeness of recorded data at item level for the years 2022 and 2023. This combination of methods yielded information about substantive and practical changes to the Ritax data over time. It also revealed patterns of missing data and variations in data recording between domains and subgroups. In addition, two focus groups were set up to discuss a data governance issue, namely whether Ritax data, if deemed sufficiently useful, could be made systematically available to the Research and Data Centre (WODC) at micro level and, if so, what would be necessary for this purpose.

The results of the qualitative analysis show that the consistency of Ritax registrations over time appears to be limited. Substantive changes to the instrument, such as the introduction of Ritax 2.0 in 2022, have resulted in a methodological break. This makes it difficult to compare data from before and after this date. Data from the date of introduction of Ritax 2.0 onwards might possibly be suitable for investigating changes over time. However, the interviews and focus groups revealed that in practice there are regional differences in the methods used to collect the Ritax data, and it is not clear whether these differences are consistent over time. This regional variation in the way Ritax data are recorded and the changes in the data over time limit its usefulness for the purpose of researching comparisons between regions and subgroups and changes in the characteristics of juvenile suspects over time. Moreover, the quantitative analyses of the subgroups revealed that the extent to which values were missing for some items differed significantly by gender, age and collection region.

A finding based on the quantitative analyses was that the Ritax data were not always fully entered, given the number of 'unknown' scores. Given the relatively large number of 'unknown' scores within the domains of family, attitudes and relationships, the extent to which the risk and protective factors within these domains are useful for research purposes is debatable. Although the proportion of missing values was slightly lower in 2023 than in 2022, consideration must be given to the possible selectivity in the Ritax data. In view of the differences in missing values for gender, age and region, the data are incomplete. This affects the representativeness of the data and limits the possibility of making comparisons over time. Data from one year or region might mostly relate to a specific subpopulation, while data from another year or region might focus on a different one. Possible shifts of this kind in the underlying population make it more difficult to interpret trends. Such selectivity can be assessed only by linking Ritax data at micro level to an external source comprising the entire population of juvenile suspects (or juvenile offenders).

Moreover, an analysis of demographic characteristics showed that age, gender, gravity of offence and regional distribution remained largely unchanged over the years 2022 and 2023. The majority of the data concern youths (approximately 90%), with a mean age of 15.4 years. As only two years were compared and there were a significant number of missing values, few conclusions can be drawn about changes in the gravity of offences.

As regards the data governance issue, it was apparent from the focus groups that, legally and technically, there are ways in which Ritax data could be provided systematically at micro level to the Research and Data Centre (WODC). As yet, there is no ongoing, substantive requirement under the General Data Protection Regulation (GDPR) that would warrant a systematic exchange of such data at micro level. The systematic provision of data would also require significant personnel and material capacity, security measures and clear agreements between the parties involved. At present, there is insufficient need for this.

This feasibility study has a few limitations and also offers a basis for follow-up. As the empirical analyses cover only two recent years, they shed no light on changes over the longer term. Moreover, as it remains unclear how the selection of the young people about whom Ritax data were recorded took place, the extent to which the data on juvenile suspects are representative is uncertain. As this study has been unable to clarify why a significant proportion of the variables are scored as 'unknown', the causes of these missing values are unclear. However, qualitative findings point to

variations in the collection of data and difficulty in obtaining information for certain items. Nevertheless, this study offers a valuable basis for further research. For the first time, there is information about the extent to which the Ritax data are consistent and complete at the national level and where differences exist between domains, regions and subgroups. If the reliability and hence the usability of the Ritax data improves in the future, these data can contribute to evaluation research, subgroup comparisons and interpretation of trends in juvenile delinquency.

All things considered, it is apparent from this feasibility study that at present there is not sufficient justification for using Ritax data *systematically* for periodic research into changes over time in the background characteristics of the total population of juvenile suspects and regional and subgroup variation within that population. In other words, it would not currently be advisable to use these data for biennial monitoring such as in the JCM. At this stage, there is too much variation in the manner in which data are recorded and entered to make reliable comparisons. If their quality is improved, Ritax data could eventually become a reliable source for such research purposes. To improve the quality and usability of Ritax data, it is particularly important to reduce the number of 'unknown' scores in domains with many missing values. Consideration could be given to achieving greater uniformity in data collection and recording procedures in practice. Naturally, allowance must be made for the fact that child welfare investigators can exercise a degree of discretion that enables them to tailor individual investigations as needed. Moreover, systematic documentation of substantive changes to the instrument is essential in order to identify potential future methodological breaks. Training of child welfare investigators and stricter national coordination of data recording guidelines could help to generate more consistent and reliable data capable of being used in investigating changes in the population of juvenile suspects over time.

The use of multi-year Ritax data for ad hoc research into changes in the background characteristics of juvenile suspects or offenders over time would seem feasible, provided attention is paid to the limitations described in this study. However, careful consideration should be given in the design of the research to how to take account of the extent of missing values and how this affects the conclusions that may be drawn from the data. In order to limit the selectivity of the information and to be able to make statements about the population of juvenile suspects, a link is needed at micro level to an external source in which data about the entire population of juvenile suspects or offenders are available (for example, the police records at Statistics Netherlands (CBS) or the Research and Policy Database for Judicial Documentation (OBJD) at the Research and Data Centre (WODC).

Literatuur

- Buyse, W., Petersen, A., & Nauta, O. (2022). *Toeleiding naar gedragsinterventies in de jeugdstrafrechtketen: Het landelijk instrumentarium en de persoonsgerichte aanpak*. DSP-Groep.
repository.wodc.nl/bitstream/handle/20.500.12832/3222/3205-toeleiding-naar-gedragsinterventies-jeugdstrafrechtketen-volledige-tekst.pdf?sequence=1&isAllowed=y
- Buyse, W., & Petersen, A. (2023). *Vernieuwde risicotaxatie Ritax 2.0. Evaluatie van implementatie en gebruik door RVDK en jeugdreclassering*. DSP-Groep.
repository.wodc.nl/bitstream/handle/20.500.12832/3300/3360-vernieuwde-risicotaxatie-ritax-volledige-tekst.pdf?sequence=1&isAllowed=y
- Directoraat-Generaal Straffen en Beschermen (2022). *20220125 Memo Monitoring Ritax 2.0* [Intern document]. Ministerie van Justitie en Veiligheid.
- Directoraat-Generaal Straffen en Beschermen Ketenregie. (2024a). *1.LOGBOEK AANPASSINGEN INSTRUMENT; Ritax* [Excel-bestand]. Ministerie van Justitie en Veiligheid.
- Egberink, I.J.L., & Meijer, R.R. (2011). *Voorstudie Convergente Validiteit LIJ*. Rijksuniversiteit Groningen.
repository.wodc.nl/bitstream/handle/20.500.12832/1937/2110-volledige-tekst_tcm28-72109.pdf?sequence=2&isAllowed=y
- Enders, C.K. (2022). *Applied missing data analysis* (2^e ed.). The Guilford Press.
www.guilford.com/books/Applied-Missing-Data-Analysis/Craig-Enders/9781462549863
- Kamerstukken II 2022/23, 28 741, nr. 107 (2023, 4 september). Overheid.nl.
[Kamerstukken II - 28741-107](https://www.rijksoverheid.nl/documenten/vergaderingen/kamerstukken/2023/09/04/kamerstukken-ii-28741-107)
- Mensink, K., Hill, J., & Weijters, G. (2021). *Op zoek naar profielen van jeugdige verdachten: Een verkenning op basis van risico- en beschermende factoren uit het Landelijk Instrumentarium Jeugdstrafrechtketen*. WODC.
repository.wodc.nl/bitstream/handle/20.500.12832/3061/Cahier-2021-11-volledige-tekst.pdf?sequence=1&isAllowed=y
- Ministerie van Justitie en Veiligheid (2011). *Handleiding Landelijk Instrumentarium Jeugdstrafrechtketen, versie Mei 2011* [Intern document]. Raad voor de Kinderbescherming.
www.rijksoverheid.nl/documenten/richtlijnen/2025/01/15/handleiding-landelijk-instrumentarium-jeugdstrafrechtketen-lij
- Ministerie van Justitie en Veiligheid (2012). *Landelijk Instrumentarium Jeugdstrafrechtketen (LIJ) – Leidraad in de keten*. Geraadpleegd op 20 Oktober, 2024, van wegwijzerjeugdenveiligheid.nl/documenten/leidraad-in-de-keten-landelijk-instrumentarium-jeugdstrafrechtketen
- Ministerie van Justitie en Veiligheid (2022). *Ritax 2.0 Puntentelling en DRP scores LIJ – DEF* [Intern document]. Ministerie van Justitie en Veiligheid.
- Ministerie van Justitie en Veiligheid (2024). *Handleiding Landelijk Instrumentarium Jeugdstrafrechtketen, versie Maart 2024* [Intern document]. Raad voor de Kinderbescherming. Zie ook www.rijksoverheid.nl/documenten/richtlijnen/2025/01/15/handleiding-landelijk-instrumentarium-jeugdstrafrechtketen-lij
- Nauta, O., & Loef, L. (2010). *Evaluatie pilot landelijk instrumentarium jeugdstrafrechtketen. Situatie tot 30 April 2010*. DSP-Groep.
repository.wodc.nl/handle/20.500.12832/1738

- Nauta, O., & Loef, L. (2011). *Evaluatie pilot landelijk instrumentarium jeugdstrafrechtketen*. DSP-Groep. www.dsp-groep.nl/wp-content/uploads/12onlandi2_Evaluatie_landelijk_instrumentarium_jeugdstrafrechtketen_Samenvatting.pdf
- Raad voor de Kinderbescherming (2024). *20240131_Gebruik_LIJ_2019 tot 2024_naarRegio* [Intern document]. Ministerie van Justitie en Veiligheid.
- Raad voor de Kinderbescherming (2018). *Gebruik LIJ in Strafonderszoeken* [Interne memorandum]. Ministerie van Justitie en Veiligheid.
- Timmermans, M., & Witvliet, M. (2011). *Screenen en signaleren in de jeugdstrafrechtketen. Een onderzoek naar de validiteit van psychisch disfunctioneren en de betrouwbaarheid van het Landelijk Instrumentarium Jeugdstrafrechtketen (LIJ)*. Regioplan Beleidsonderzoek. repository.wodc.nl/bitstream/handle/20.500.12832/1761/volledige-tekst_tcm28-70525.pdf?sequence=2&isAllowed=y
- Tollenaar, N., Beijers, J., & Van der Laan, A.M. (2024). *Monitor zelfgerapporteerde jeugddelinquentie 2023*. WODC. <https://repository.wodc.nl/handle/20.500.12832/3385>
- Van der Ark, L.A., Van Leeuwen, J.L., & Jorgensen, T.D. (2018). *Interbeoordelaarsbetrouwbaarheid LIJ: Onderzoek naar de interbeoordelaarsbetrouwbaarheid van het Landelijk Instrumentarium Jeugdstrafrechtketen*. Universiteit van Amsterdam. repository.wodc.nl/handle/20.500.12832/2267
- Van der Laan, A.M., Tollenaar, N., Beijers, J., & Kessels, R.J. (2024). *Ontwikkelingen in de jeugdcriminaliteit 2000-2023. Synthese van bevindingen uit de Monitor Jeugdcriminaliteit*. WODC. repository.wodc.nl/handle/20.500.12832/3384
- Van der Put, C., & Kapteins, L. (2020). *Normeringonderzoek instrumenten 2A en 2B van het Landelijk Instrumentarium Jeugdstrafrecht (LIJ)*. Universiteit van Amsterdam. hdl.handle.net/11245.1/fa324d85-b079-4277-886a-2248720974db
- Van der Put, C. & Stolwijk, I.J. (2022). *Normering en actualisering van het Halt-signaleringsinstrument (Halt-SI) en beschrijving van de actuele populatie Halt-deelnemers*. Universiteit van Amsterdam. <https://repository.wodc.nl/handle/20.500.12832/3186>
- Van der Put, C., Spanjaard, H., Van Domburgh, L. et al. (2011). Ontwikkeling van het Landelijke Instrumentarium Jeugdstrafrechtketen (LIJ), *Kind & Adolescent Praktijk*, 10 (1), 76–83. <https://doi.org/10.1007/s12454-011-0021-2>
- WODC. (n.d.). Monitor Jeugdcriminaliteit. Geraadpleegd op 3 februari 2025, van www.wodc.nl/onderwerpen/monitor-jeugdcriminaliteit

Bijlage 1 Aantallen en percentages van ontbrekende waarden op alle items in de dataset van 2023 en 2022

Online downloadbare tabel: [Aantallen en percentages van ontbrekende waarden op alle items in de datasets van 2023 en 2022](#)

Bijlage 2 Alle items met een significant verband voor het jaar 2023 en 2022 op basis van sekse, leeftijd, geslacht en afnameregio

Online downloadbare tabel: [Alle items met een significant verband voor het jaar 2023 en 2022 op basis van sekse, leeftijd, geslacht en afnameregio](#)

Het Wetenschappelijk Onderzoek- en Datacentrum (WODC), Kennisinstituut voor de rechtsstaat, is een onafhankelijk kennisinstituut dat valt onder het ministerie van Justitie en Veiligheid. Het WODC draagt bij aan behoud en verbetering van de rechtsstaat via het (laten) uitvoeren van kwalitatief hoog wetenschappelijk onderzoek. En door het aanbieden van gevraagde en ongevraagde kennis, verbeterpunten en (waar mogelijk) denkrichtingen.

Voor meer informatie, bezoek www.wodc.nl