



Toepassing van anonieme onderzoeksgegevens over de SDQ binnen de JGZ

Bruikbaarheid en valkuilen

H. Hoofs · N. Jansen · D. Mohren · M. Jansen · IJ. Kant

Published online: 14 March 2018

© Bohn Stafleu van Loghum is een imprint van Springer Media B.V., onderdeel van Springer Nature 2018

Samenvatting Om psychosociale problemen binnen de jeugdgezondheidszorg (JGZ) tijdig te signaleren bij basis- en middelbare scholieren wordt er tijdens de periodieke gezondheidsonderzoeken (PGO) gebruik gemaakt van de Strengths and Difficulties Questionnaire (SDQ). Voor de ondersteuning van deze vroegsignalering met behulp van de SDQ zijn er afkappunten opgesteld. Deze afkappunten komen voornamelijk voort uit validatiestudies met een anonieme setting. Deze setting verschilt met de vertrouwelijke, niet-anonieme setting van het PGO, waar de professional inzicht krijgt in de vragenlijst. Hierdoor is er dus een discrepantie tussen de setting waarin de SDQ wordt toegepast en waar deze wordt gevalideerd. In deze studie wordt onderzocht of deze discrepantie zorgt voor een verschil in de SDQ met betrekking tot de betrouwbaarheid, gemiddelde scores en afkappunten. Hiervoor worden gegevens gebruikt van de vragenlijsten uit het vertrouwelijke PGO ($N=6.594$) en

de anonieme Jeugdmonitor ($N=4.613$), beide afgenomen in het tweede jaar van het voortgezet onderwijs. De betrouwbaarheid in beide settings is gelijk; de gemiddelde scores en afkappunten verschillen echter significant. Het is dus niet valide om de afkappunten uit de anonieme validatie studies een-op-een te gebruiken in het PGO, dan wel gemiddelde scores tussen deze settings direct met elkaar te vergelijken.

Trefwoorden jeugdgezondheidszorg · SDQ · vroegsignalering · afkappunten · psychosociale problemen

Inleiding

De jeugdgezondheidszorg (JGZ) volgt de lichamelijke, psychische, sociale en cognitieve ontwikkeling van kinderen in Nederland van 0 tot 19 jaar en interverteert waar nodig om ontwikkelingsstoornissen en de gevolgen daarvan te voorkomen en te beperken [1]. Aangezien ongeveer een op de tien basisscholieren in Nederland psychosociale problemen heeft [2], en de gevolgen hiervan een grote impact hebben op zowel het individu als de maatschappij, is een van de kerntaken van de JGZ het tijdig signaleren en opsporen van psychosociale problemen [3]. Voor een tijdige signalering van psychosociaal probleemgedrag in de JGZ, wordt er gebruik gemaakt van de Nederlandse versie van de Strengths and Difficulties Questionnaire (SDQ) [4–7]. Deze vragenlijst bestaat uit 25 items die kunnen worden onderverdeeld in vijf subschalen (emotionele symptomen, gedragsproblemen, hyperactiviteit of aandachtstekort, problemen met leeftijdgenoten en prosociaal gedrag). Daarnaast kan er een totaalscore worden berekend waarbij de subschaal prosociaal gedrag niet meegenomen wordt. Doordat de SDQ bewezen heeft goed te functioneren als instrument voor vroegsignalering, het een compacte vragenlijst is, en er versies zijn voor zowel

Dit artikel is eerder gepubliceerd in: Tijdschr Gezondheidswet 2016;5:181–88 en is een bewerking door de auteurs van hun onderzoek dat in 2015 verscheen in het tijdschrift PLoS One (10(4): e0120930).

H. Hoofs (✉) · N. Jansen · IJ. Kant
Vakgroep Epidemiologie, CAPHRI, Universiteit Maastricht,
Maastricht, Nederland
huub.hoofs@maastrichtuniversity.nl

H. Hoofs · M. Jansen
Academische Werkplaats, GGD Zuid Limburg, Maastricht,
Nederland

D. Mohren
MEMIC Maastricht, Centrum voor data- en
informatiemanagement, Universiteit Maastricht,
Maastricht, Nederland

M. Jansen
Vakgroep Health Services Research, CAPHRI, Universiteit
Maastricht, Maastricht, Nederland



Noot van de redactie

In 2016 is de SDQ handleiding aangepast [17]. Hierin zijn Nederlandse afkappunten voor de SDQ opgenomen. De data die gebruikt zijn bij het bepalen van deze afkappunten voor de ouder- én leerkrachtversie van de SDQ zijn verzameld in het kader van periodieke gezondheidsonderzoeken (PGO). De twee uitzonderingen hierop zijn de ouder- én leerkrachtversie 4- tot 7-jarigen. Bij het bepalen van de afkappunten voor deze twee versies zijn twee databronnen gebruikt. Een databron met data die verzameld zijn in een anonieme setting en de andere databron met data die verzameld zijn in het kader van een PGO.

ouders, kinderen (jeugdversie) als leerkrachten wordt deze vragenlijst landelijk gebruikt binnen de JGZ [8]. Geadviseerd wordt de vragenlijst tijdens drie periodieke gezondheidsonderzoeken (PGO) af te nemen, dat wil zeggen in de onderbouw en bovenbouw van het basisonderwijs (4 tot 6 jaar en 9 tot 11 jaar) en in het voortgezet onderwijs (12 tot 16 jaar) [9]. In de onder- en bovenbouw van het basisonderwijs, wordt de vragenlijst door de ouder(s) of voogd van het kind ingevuld en tijdens het voortgezet onderwijs door het kind zelf. De SDQ heeft naast het primaire doel van tijdige signalering van psychosociaal probleemgedrag bij kinderen, als secundair doel de monitoring van de psychosociale ontwikkeling van de jeugd in Nederland. Zodoende is de SDQ een van de indicatoren van de Inspectie voor de Gezondheidszorg (IGZ) [10] en maakt het onderdeel uit van het project Jeugd in Beeld waarmee de implicaties voor zorg, onderzoek, en beleid worden vastgesteld [11].

Bij de toepassing van de SDQ wordt gebruik gemaakt van afkappunten. Bij scores boven een bepaald afkappunt is er een verhoogde kans dat een kind zich niet optimaal ontwikkelt [12]. Een score boven het afkappunt helpt dan ook, in combinatie met het consult, bij de objectivering van de oordeelsvorming over eventuele psychosociale problemen of risico's bij een kind [13]. Als blijkt dat het kind inderdaad een verhoogde kans heeft op eventuele problemen of risico's kan de professional het kind uitnodigen voor een vervolgesprek of doorverwijzen naar de eerste- of tweedelijnszorg. Afkappunten voor de SDQ zijn voornamelijk gebaseerd op de tachtigste en de negentigste percentiel scores. Hierbij gaat het dus om de kinderen met de hoogste score binnen een bepaalde groep. Deze percentielscores komen voort uit het oorspronkelijke werk van Robert Goodman et al. [4, 14–16] die vonden dat deze scores de onderzochte populatie het best onderverdeelden in kinderen met een normaal en verhoogd risico voor psychosociale aandoeningen uit het diagnostisch en statistisch handboek voor psychische stoornissen IV (DSM-IV; bijvoorbeeld ADHD). Bij de herevaluatie van de SDQ in de Neder-

landse populatie zijn deze percentielscores opnieuw berekend voor verschillende doelgroepen in een anonieme setting [5]. Deze afkappunten worden thans toegepast binnen de JGZ om zodoende de oordeelsvorming te faciliteren (zie kader met noot van de redactie). De (anonieme) setting waarin deze afkappunten oorspronkelijk werden gedefinieerd verschilt echter van de setting waar deze afkappunten worden toegepast, het PGO. In dit artikel wordt gekeken of dit verschil in setting resulteert in een incorrect gebruik van afkappunten wanneer deze in een andere setting worden toegepast en wat de gevolgen hiervan zijn.

De SDQ wordt dus in het algemeen gevalideerd in een anonieme setting. In een anonieme setting wordt de SDQ bijvoorbeeld klassikaal afgenomen zonder vermelding van naam of geboortedatum. Hierdoor zal er minder druk van buitenaf worden ervaren om sociaal wenselijke antwoorden te geven. Binnen een PGO wordt de vragenlijst echter vertrouwelijk ingevuld, omdat deze wordt besproken gedurende het consult. Hierdoor kan het kind mogelijk een bepaalde druk ervaren om sociaal wenselijke antwoorden te geven, hetgeen kan zorgen voor *selfreport bias*. *Self-report bias* verwijst naar de vertekening die kan optreden als mensen informatie over zichzelf verstrekken door middel van een vragenlijst. Deze bias kan versterkt worden omdat een verhoogde score kan leiden tot een follow-up of doorverwijzing. Als dit gevolg bekend is bij het kind, en als ongewenst wordt ervaren vanwege mogelijke vervolgacties, kan het kind, al dan niet bewust, meer terughoudend zijn bij het beantwoorden van de vragenlijst. Dit kan resulteren in een onderrapportage van het probleemgedrag. Dergelijke onderrapportage is voor de SDQ nog nooit onderzocht, maar wel al meermaals aangetoond bij andere vragenlijsten [18]. Deze studies laten zien dat met betrekking tot het effect van sociaal wenselijkheid meisjes gevoeliger zijn voor het gebrek aan anonimiteit dan jongens en dus sterker geneigd zijn tot onderrapportage [19].

Met betrekking tot de betrouwbaarheid van de SDQ lijkt de betrouwbaarheid van de totale probleemscore goed, maar van de vijf subschalen is deze over het algemeen aan de lage kant [20–23]. De theoretische onderbouwing van de subschalen is echter meer dan voldoende en ook de construct-, indruks- en inhoudsvaliditeit zijn goed [24, 25]. Omdat de betrouwbaarheid eveneens zou kunnen afhangen van de setting waarin de vragenlijst wordt afgenomen zal dit ook worden geanalyseerd in het huidige artikel.

Om dit alles te onderzoeken wordt gebruik gemaakt van data verzameld binnen de JGZ in de PGO's (vertrouwelijke setting) en in de Jeugdmonitor (anonieme setting). De Jeugdmonitor is een anonieme monitor die is bedoeld om beleidsmakers, onderzoekers en andere geïnteresseerden te informeren over de situatie van de jeugd. Door gebruik te maken van data van deze twee settings is de volgende onderzoeksvraag beantwoord: In welke mate is de jeugdversie van de SDQ

afgenomen in het PGO vergelijkbaar met deze afgenomen in de Jeugdmonitor, betreffende de betrouwbaarheid, gemiddelden, en afkappunten?

Methode

Data

Voor de analyse omtrent de vertrouwelijke afname van de SDQ is gebruik gemaakt van een reeds lopende dataverzameling bij de gemeentelijke gezondheidsdienst (GGD) Zuid Limburg (Maastricht, Heerlen, Sittard-Geleen en de vijftien omliggende buurgemeenten). Voorafgaand aan het PGO krijgt ieder kind een vragenlijst thuisgestuurd. Om een breed beeld te krijgen van het welzijn van het kind bevat de PGO-vragenlijst naast de SDQ ook andere vragen. Er wordt gevraagd aan het kind deze vragenlijst in te vullen en mee te nemen naar het PGO. Voor de huidige studie zijn de vragenlijsten gebruikt van de PGO's van de tweede klas van het voortgezet onderwijs. Deze PGO's bestrijken het gehele schooljaar (mediaan=maart). De vertrouwelijke data zijn geanonimiseerd. Data van het schooljaar 2010–2011 en 2011–2012 zijn gebruikt. Gezien de zorgplicht van de JGZ kent het PGO een opt-out procedure waarbij het kind dan wel de ouder(s) of voogd actief moeten aangeven als ze niet willen deelnemen. In de informatiebrief voorafgaand aan het onderzoek wordt duidelijk dat de vragenlijst door een jeugdverpleegkundige wordt ingezien. De gegevens van ongeveer 90% van alle kinderen die in aanmerking komen voor het PGO in de tweede klas van het voortgezet onderwijs zijn aanwezig [26]. Dit resulteerde in een totaal van 6.594 bruikbare vragenlijsten. De gemiddelde leeftijd van de kinderen was 14,26 jaar met een standaard deviatie (SD) van 0,62 jaar. Hiervan was 51,30% meisje.

Voor de anonieme setting is gebruik gemaakt van de resultaten van de Jeugdmonitor in Zuid-Limburg. Deze monitor wordt uitgezet om een indicatie te krijgen van het welzijn van de jeugd in Zuid-Limburg en bevat onder meer de SDQ ter indicatie van psychosociale problematiek voor adolescenten in de tweede klas van het voortgezet onderwijs. De Jeugdmonitor bestaat uit een elektronische vragenlijst die, geheel anoniem, tijdens school wordt ingevuld. In sommige gevallen is om technische redenen (17%) gebruik gemaakt van een papieren vragenlijst. Gegevens van de Jeugdmonitor afgenomen in het laatste kwartaal van 2009 (mediaan= november) zijn gebruikt. Aangezien het hier om regulier vragenlijst onderzoek gaat wordt er gebruik gemaakt van een opt-in procedure, waarbij de ouder(s)/voogd toestemming moet(en) geven voor deelname. In de informatiebrief wordt de anonimiteit van het onderzoek duidelijk vermeld. Ongeveer 70% van de uitgenodigde kinderen hebben deelgenomen. Dit resulteerde in een totaal van 4.613 vragenlijsten. De gemiddelde leeftijd van de respondenten was 13,80 jaar (SD=0,57 jaar). Hiervan was 50,01% meisje.

De populaties van de vertrouwelijke PGO-gegevens en die van de anonieme Jeugdmonitorgegevens zijn vergelijkbaar: Het betreft beide leerlingen in de tweede klas van het regulier middelbaar onderwijs in Zuid-Limburg. Met als verschil de mate van anonimiteit waarin de vragenlijsten worden afgenomen.

Instrumenten

SDQ De primaire uitkomstmaat is de SDQ. De vijf subschalen (emotionele symptomen, gedragsproblemen, hyperactiviteit of aandachtstekort, problemen met leeftijdgenoten, en prosociaal gedrag) lopen van 0 tot 10. Daarnaast kan er op basis van de subschalen een totaalscore worden berekend, lopend van 0 tot 40, waarbij de subschaal prosociaal gedrag niet meegenomen wordt. De 25 stellingen (bijvoorbeeld: *Ik pieker veel*) kunnen allen op een driepuntenschaal worden gescoord (*niet waar, een beetje waar, zeker waar*). Een vijftal items zijn gehercodeerd bij het berekenen van de somscores.

Achtergrondvariabelen De achtergrondvariabelen leeftijd, geslacht en thuissituatie zijn direct onttrokken uit de vragenlijsten. Stedelijkheid van woonomgeving en sociaal economische status (SES) zijn gedefinieerd op basis van de 4-cijferige postcodes van de respondenten [27]. Respondenten die woonden in een omgeving met gemiddeld meer dan 1.000 omgevingsadressen per vierkante km zijn geclassificeerd als wonend in een stedelijk gebied tegenover wonend in een landelijk gebied voor respondenten met minder dan 1.000 omgevingsadressen per vierkante km [27]. SES is een relatieve schaal, berekend op basis van verschillende gegevens zoals WOZ-waarde en fiscaal maandinkomen, waarbij een hogere score is gerelateerd aan een hogere SES [28]. Om de respondenten te classificeren als autochtoon of allochtoon (westers of niet-westers) is er gebruik gemaakt van de CBS-definitie [29].

Analyse

De achtergrondkenmerken van de steekproeven uit beide settings zijn vergeleken. Om de betrouwbaarheid te berekenen is gebruik gemaakt van de Cronbach's α . Hierbij worden de volgende richtlijnen gehanteerd: Uitstekend ($\alpha > 0,9$), goed ($0,7 < \alpha < 0,9$), acceptabel ($0,6 < \alpha < 0,7$), slecht ($0,5 < \alpha < 0,6$), en onacceptabel ($\alpha < 0,5$). Omdat de bruikbaarheid van de Cronbach's α ter discussie staat, aangezien het geen goede maat zou zijn voor de interne consistentie en de betrouwbaarheid kan onderschatten, is ook gebruik gemaakt van de grootste ondergrens (*greatest lower bound (glb)*) [30]. Deze glb is een betere benadering van de daadwerkelijke interne consistentie. Aangezien de glb altijd groter is dan de Cronbach's α zullen beide maten als een interval worden weergegeven. De gemiddelde scores op de subschalen en de totale score van de SDQ binnen de Jeugdmonitor en

Tabel 1 Achtergrondkenmerken voor de vertrouwelijke PGO setting en de anonieme Jeugdmonitor setting

	PGO	Jeugdmonitor	verschil test
aantal (N)	6.422	4.560	
<i>geslacht (%)</i>			
jongen	48,72	49,78	$\chi^2(1) = 1,15, p = 0,28$
meisje	51,28	50,22	
<i>leeftijd</i>			
gemiddelde (SD ^a)	14,26 (0,62)	13,81 (0,57)	$t(10.978) = 39,36, p < 0,01$
<i>etniciteit (%)</i>			
autochtoon	84,56	81,51	$\chi^2(2) = 18,69, p < 0,01$
allochtoon (westers)	8,54	10,64	
allochtoon (niet-westers)	6,91	7,86	
<i>sociaal economische status</i>			
gemiddelde (SD ^a)	-0,33 (1,13)	-0,49 (1,13)	$t(10.953) = 7,40, p < 0,01$
<i>thuisituatie (%)</i>			
2 biologische ouders	77,95	75,48	$\chi^2(2) = 66,86, p < 0,01$
1 biologische ouder	20,74	20,84	
anders/onbekend	1,31	3,68	
<i>woongebied (%)</i>			
landelijk	51,70	47,66	$\chi^2(1) = 13,74, p < 0,01$
stedelijk	48,30	52,34	
^a SD standaard deviatie			

het PGO zijn met elkaar vergeleken door middel van een variantieanalyse ($\alpha = 0,05$). Hierbij is er ook onderscheid gemaakt naar geslacht. Bij een significant interactie effect is er gebruik gemaakt van een post-hoc toets (Tukey) om te toetsen welke groepen van elkaar verschillen [31]. Als effectmaat is de η^2 gebruikt. De η^2 is een indicatie van de door de onafhankelijke variabele verklaarde variantie in de afhankelijke variabele (bijvoorbeeld totale probleemscore). De volgende richtlijnen zijn gebruikt voor interpretatie van de η^2 : klein (0,01), middel (0,06), en groot (0,16) [32].

De gemiddelde scores op de subschalen en de totale score van de SDQ binnen de Jeugdmonitor en het PGO zijn met elkaar vergeleken door middel van een variantieanalyse ($\alpha = 0,05$). Hierbij is er ook onderscheid gemaakt naar geslacht. Bij een significant interactie effect is er gebruik gemaakt van een post-hoc toets (Tukey) om te toetsen welke groepen van elkaar verschillen [31]. Als effect maat is de η^2 gebruikt. De Z2 is een indicatie van de door de onafhankelijke variabele verklaarde variantie in de afhankelijke variabele (bijvoorbeeld totale probleemscore). De volgende richtlijnen zijn gebruikt voor interpretatie van de η^2 : klein (0,01), middel (0,06), en groot (0,16) [32].

Daarnaast zijn de afkappunten berekend voor de gegevens van zowel de Jeugdmonitor als het PGO. Op deze manier kan het percentage kinderen dat een 'normale', 'grens', of 'afwijkende' score heeft in beide settings berekend worden [5]. Normale scores zijn gebaseerd op het tachtigste percentiel in de Jeugdmonitor. Grensscores lopen van dit percentiel tot en met het negentigste percentiel en alle scores hoger dan het negentigste percentiel zijn als afwijkend gedefinieerd. Deze onderverdeling komt overeen met

de oorspronkelijke validatie binnen de Nederlandse populatie [5].

Resultaten

Kenmerken van de steekproef

De verdeling van jongens en meisjes is in beide settings gelijk (zie tab. 1). Er zijn enkele verschillen tussen de twee groepen. Zo zijn de respondenten uit het PGO ouder dan de respondenten uit de Jeugdmonitor. Er zijn binnen de Jeugdmonitor minder autochtonen en meer westerse allochtonen in vergelijking met het PGO. Ook het aantal respondenten met twee biologische ouders ligt lager in de Jeugdmonitor ten opzichte van het PGO. Verder wonen respondenten in de Jeugdmonitor vaker in een stedelijk gebied en hebben ze gemiddeld een lagere SES.

Betrouwbaarheid van de SDQ

Zowel de totale probleemscore (Cronbach's α -glb = 0,78–0,85) als de hyperactiviteit of aandachtstekort subschaal (0,76–0,81) lieten een hoge betrouwbaarheid zien. De subschalen prosociaal gedrag (0,65–0,67) en emotionele symptomen (0,69–0,71) lieten een gemiddelde betrouwbaarheid zien, terwijl de betrouwbaarheid voor de gedragsproblemen (0,55–0,56) en problemen met leeftijdgenoten (0,51–0,55) laag waren. De betrouwbaarheid is ook per setting apart berekend. In het algemeen is de betrouwbaarheid gelijk aan de hiervoor gerapporteerde betrouwbaarheidscoëfficiënten waarin de twee settings samengevoegd zijn. De betrouwbaarheid van de totale probleemscore is

Tabel 2 Vergelijking, door middel van een variantie-analyse, van de totale probleemscore en de subschalen van de SDQ voor setting en geslacht

	setting; gemiddelde ^a (standaard deviatie)		η^2 (effectgrootte) ^b		
	PGO	Jeugdmonitor	setting	geslacht	interactie
<i>totale probleemscore</i>			0,04*	0,01*	0,002*
jongens	7,31 ^f (4,49)	8,96 ^f (5,12)			
meisjes	7,81 ^g (4,83)	10,27 ^h (5,20)			
<i>emotionele symptomen</i>			0,02*	0,09*	0,004*
jongens	1,42 ^f (1,54)	1,66 ^f (1,77)			
meisjes	2,40 ^g (2,05)	3,12 ^h (2,29)			
<i>gedragsproblemen</i>			0,05*	0,005*	<0,001
jongens	1,27 (1,26)	1,89 (1,65)			
meisjes	1,09 (1,14)	1,70 (1,44)			
<i>hyperactiviteit/aandacht tekort</i>			0,01*	<0,001	0,001*
jongens	3,58 ^f (2,46)	3,93 ^f (2,43)			
meisjes	3,35 ^g (2,38)	4,08 ^h (2,45)			
<i>problemen met leeftijdgenoten</i>			0,02*	0,001*	<0,001
jongens	1,04 (1,33)	1,47 (1,60)			
meisjes	0,97 (1,30)	1,37 (1,48)			
<i>prosociaal gedrag</i>			0,05*	0,06*	0,001*
jongens	7,92 ^f (1,62)	7,02 ^f (2,00)			
meisjes	8,67 ^g (1,34)	7,95 ^h (1,73)			

*Significant effect ($\alpha = 0,05$)
^aAls er een significant interactie effect is geven superscripts aan welke groepen van elkaar verschillen, waarbij verschillende superscripts een significant verschil aangeven
^bEffectgrootte (η^2) is de proportie verklaarde variantie door de onafhankelijke variabele in de ANOVA, met de volgende richtlijn: klein (0,01), middel (0,06), en groot (0,14)

zowel voor het PGO (Cronbach's α -glb = 0,77–0,85) als voor de Jeugdmonitor (Cronbach's α -glb = 0,76–0,85) dan ook goed.

Vergelijking van SDQ-scores tussen setting en geslacht

Om te kijken of er structurele verschillen zijn in de gemiddelde scores van de totale probleemscore en subschalen is een variantieanalyse uitgevoerd. Er is een significant verschil tussen het PGO en de Jeugdmonitor voor alle schalen (zie tab. 2) gevonden. Gemiddelden zijn hoger voor de Jeugdmonitor vergeleken met het PGO, behalve voor de prosociaal gedrag subschaal. Deze subschaal is echter een index voor positief gedrag waardoor de richting van de associatie hetzelfde is. Geslacht is significant geassocieerd met alle schalen behalve de hyperactiviteit of aandachtsstekort subschaal. Interactie-effecten zijn significant voor de subschalen emotionele symptomen, hyperactiviteit of aandachtsstekort en prosociaal gedrag, en voor de totale probleemscore. Tab. 2 laat zien dat als er een significant interactie-effect is, meisjes een groter verschil laten zien tussen scores van het PGO en de Jeugdmonitor in vergelijking met jongens.

Afkappunten van de SDQ

De afkappunten gebaseerd op het tachtigste en negentigste percentiel zijn lager in het PGO vergeleken

met die van de Jeugdmonitor (tab. 3). De verschillen tussen jongens en meisjes zijn overeenkomstig de verschillen in gemiddelden. Er zijn bijvoorbeeld meer meisjes met een grens- of afwijkende score vergeleken met jongens binnen het PGO voor de hyperactiviteit of aandachtsstekort subschaal, terwijl dit beeld omgekeerd is binnen de Jeugdmonitor. Eenzelfde beeld is te zien voor de gemiddelde scores (tab. 3). De afkappunten van de Jeugdmonitor zijn toegepast op de gegevens van het PGO. De respondenten in het PGO zijn dus gecategoriseerd op basis van de afkappunten uit de Jeugdmonitor (tab. 3). Dit resulteert in lagere prevalenties van respondenten met een grens- of afwijkende score. Op de totale probleemscore bijvoorbeeld, wordt 'slechts' 8% van de kinderen geclassificeerd als niet normaal, in plaats van de oorspronkelijke 18%. Als er alleen naar de afwijkende scores wordt gekeken dan zakt dit percentage zelfs naar 3% in plaats van de oorspronkelijke 8%.

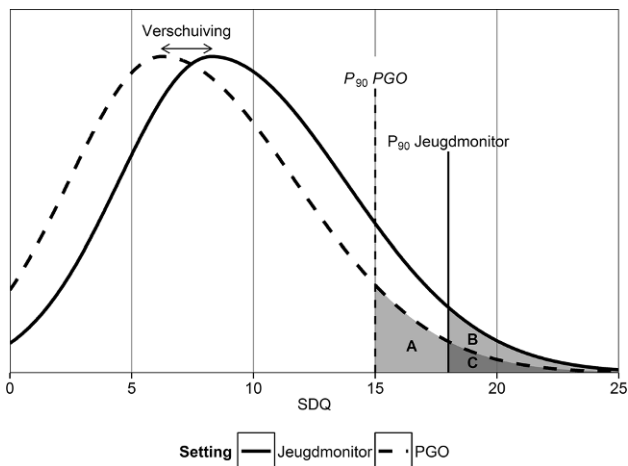
Discussie

De SDQ is een geschikt instrument voor de vroegsignalering van kinderen met een verhoogd risico op psychosociale problemen of aandoeningen. Vooral de totale probleemscore van de SDQ is een uitstekende voorspeller voor psychosociale problematiek. De toevoeging van de SDQ aan het PGO in het basisonderwijs is dan ook zeer waardevol. Vooral omdat onderzoek heeft aangetoond dat bij de

Tabel 3 Percentages voor grens- en afwijkende scores binnen het PGO en de Jeugdmonitor gebaseerd op de bijbehorende afkappunten alsmede voor het PGO op basis van de afkappunten van de Jeugdmonitor

	normaal: 0–80 percentiel			grens: 80–90 percentiel				afwijkende: >90 percentiel				
	score	T (%)	J (%)	M (%)	score	T (%)	J (%)	M (%)	score	T (%)	J (%)	M (%)
<i>Jeugdmonitor</i>												
totaal	0–14	83	86	79	15–17	9	7	11	18–40	9	7	10
emotie	0–4	83	91	74	5–6	12	7	16	7–10	6	2	10
gedrag	0–3	87	86	89	4	6	6	6	5–10	7	8	5
hyperactiviteit	0–6	84	85	83	7	6	6	7	8–10	10	9	10
leeftijdsgenoten	0–3	89	88	91	4	6	7	5	5–10	5	6	4
prosociaal	10–6	85	79	90	5	8	11	6	4–0	7	10	4
<i>PGO</i>												
totaal	0–11	82	83	80	12–14	10	10	10	15–40	8	7	10
emotie	0–3	82	90	74	4–5	13	8	18	6–10	5	2	8
gedrag	0–2	88	86	90	3	7	9	6	4–10	5	6	4
hyperactiviteit	0–5	80	79	82	6–7	13	14	12	8–10	7	8	6
leeftijdsgenoten	0–2	87	87	88	3	7	7	6	4–10	6	6	5
prosociaal	10–7	87	81	93	6	7	10	4	5–10	6	8	3
<i>PGO met afkappunten van de Jeugdmonitor</i>												
totaal	0–14	92	93	90	15–17	5	4	6	18–40	3	2	4
emotie	0–4	90	95	84	5–6	7	4	11	7–10	3	1	5
gedrag	0–3	95	94	96	4	3	3	3	5–10	2	3	1
hyperactiviteit	0–6	88	86	89	7	6	6	5	8–10	7	8	6
leeftijdsgenoten	0–3	94	94	95	4	3	3	3	5–10	2	3	2
prosociaal	10–6	94	92	97	5	4	6	2	4–0	2	3	1

T Totaal, J Jongens, M Meisjes

**Figuur 1** Theoretische verdeling voor de SDQ-scores van het PGO- en de Jeugdmonitordata. Dit illustreert de verschuiving van de verdeling tussen de twee settings. 90e Percentiel afkappunten zijn getoond voor beide settings. De grijsgekleurde gebieden illustreren de kinderen die vallen in het afwijkende bereik (~10%) voor het PGO (A+C) en de Jeugdmonitor (B+C) met behulp van de respectievelijke afkappunten. Kinderen in het PGO die in het C-gebied vallen, hebben afwijkende scores gedefinieerd door het afkappunt van de Jeugdmonitor

afwezigheid van een dergelijke objectivering van gedrag veel problemen over het hoofd worden gezien [6]. Het gebruik van afkappunten vereenvoudigt de toepassing van de SDQ doordat kinderen worden ingedeeld in twee of drie categorieën. Hierbij hebben kinderen in de hoogste categorieën (grens- en of afwijkende scores) een verhoogd risico ten opzichte van kinderen in de laagste categorie (normaal) voor psychosociale problemen. Dit onderzoek laat echter zien dat het niet mogelijk is om afkappunten gebaseerd op een anonieme setting een-op-een te gebruiken in de vertrouwelijke setting van het PGO.

Zoals geïllustreerd in fig. 1 zijn de scores in de Jeugdmonitor systematisch hoger in vergelijking met die van het PGO. Hierdoor verschuift als het ware de verdeling van de scores van het PGO naar links. In de praktijk worden de afkappunten echter bepaald op basis van anonieme studies zoals de Jeugdmonitor. Hierbinnen wordt dan het afwijkende bereik bepaald op basis van het 90^e percentiel om zodoende de 10% respondenten met de hoogste scores te selecteren (gebied B+C in fig. 1). Als dit afkappunt wordt gebruikt binnen het PGO dan resulteert dit niet in de verwachte 10% maar in een lager percentage (gebied C). Aangezien risicoschattingen worden gebaseerd op de 10% hoogste scores, hebben kinderen die alsnog boven dit 'te hoge' afkappunt scores bovendien een nog groter risico op psychosociale problemen dan verwacht, omdat bij iedere 1-puntstoename op de

SDQ het risico op psychosociale problemen groter wordt [33]. Stel dus dat een jeugdverpleegkundige in een vertrouwelijke setting als het PGO een klas van 30 kinderen ziet, dan is de verwachting dat gemiddeld genomen drie kinderen boven het afkappunt zullen scoren, maar de verpleegkundige vindt waarschijnlijk slechts een leerling die boven het op anonieme studies gebaseerde afkappunt scoort. Deze ene leerling heeft dan ook nog eens een hoger risico voor toekomstige problemen dan gedacht. De alarmbellen die afgaan bij een verhoogde score moeten dus eigenlijk nog harder afgaan dan nu het geval is. Daarnaast neemt binnen het PGO de zekerheid dat een kind die onder het op anonieme studies gebaseerde afkappunt scoort geen toekomstige problemen ontwikkelt ook sterk af. Zodoende is er sprake van een verhoogde positief voorspellende waarde en een verlaagde negatief voorspellende waarde van de SDQ als afkappunten van een anonieme setting (zoals de Jeugdmonitor) worden gebruikt binnen het PGO. Voor een correct gebruik van afkappunten van de SDQ binnen het PGO zullen deze afkappunten dus moeten worden bijgesteld op basis van gegevens uit het PGO. Hierbij kan gebruik worden gemaakt van landelijke gegevens met betrekking tot de SDQ om zodoende landelijke afkappunten te vervaardigen. Een andere optie is om de bijbehorende positief voorspellende waarde en negatief voorspellende waarde van de gebruikte afkappunten opnieuw te evalueren binnen het PGO. Uiteraard geldt dit niet alleen voor het PGO maar ook iedere andere setting waarin de SDQ op een vertrouwelijke manier wordt toegepast (zie kader: noot redactie).

Het feit dat afkappunten en gemiddelden van de SDQ niet direct vergelijkbaar zijn tussen anonieme en vertrouwelijke settings is echter niet alleen een probleem voor de vertrouwelijke settings an sich, maar ook voor het vergelijk van vertrouwelijke met anonieme settings. Zoals Mieloo en collega's aangaven in hun discussie over het gebruik van scores uit het PGO: "Er [is] voorzichtigheid geboden bij het generaliseren van onze bevindingen naar een anonieme setting" [22]. Onze studie laat zien dat een dergelijke voorzichtigheid gegrond lijkt, aangezien het een-op-een vergelijken van scores tussen een anonieme en vertrouwelijke setting een vertekend beeld geeft. Het samenvoegen van gegevens uit een anonieme en vertrouwelijke setting zoals in recent onderzoek van TNO [34] is dan ook sterk af te raden. Ook laten de resultaten zien dat de richting en grootte van verschillen tussen jongens en meisjes kunnen afhangen van de setting waarin de resultaten zijn verkregen.

De lagere scores binnen het PGO in vergelijking met de Jeugdmonitor komen waarschijnlijk voort uit de verhoogde druk om sociaal wenselijke antwoorden te geven omdat de vragenlijst afgenomen in het kader van het PGO uiteindelijk wordt ingezien door de jeugdarts dan wel jeugdverpleegkundige. Daarbij kunnen verhoogde scores zorgen voor een doorverwijzing

naar de eerste- of tweedelijnszorg of een follow-up door de jeugdarts of jeugdverpleegkundige. Hierdoor kan het kind, bewust of onbewust, een onderrapportage geven om zodoende 'buiten schot' te blijven. Er kan daarbij zowel sprake zijn van een onderrapportage van probleemgedrag als een overrapportage van positief gedrag [19]. Ook tonen de resultaten aan dat meisjes, zoals verwacht, meer worden beïnvloed door de mate van anonimiteit in vergelijking met jongens [18, 35]. Gerelateerd aan sociale wenselijkheid is het omstandereffect. Deze vorm van zelf-gerapporteerde bias verwijst naar het feit dat respondenten vragenlijsten anders invullen in aanwezigheid van andere personen (omstanders), onafhankelijk of deze omstanders de antwoorden ook daadwerkelijk inzien. Eerder onderzoek heeft aangetoond dat adolescenten in aanwezigheid van hun ouders geneigd zijn terughoudender te scoren op vragenlijsten over probleemgedrag [36]. Aangezien de vragenlijsten binnen het PGO voornamelijk thuis worden ingevuld terwijl de Jeugdmonitor op school wordt afgenomen is het niet ondenkbaar dat dit omstandereffect verschillend is voor de twee settings [37]. Dit kan nog eens verder worden versterkt door het gebruik van papieren vragenlijsten binnen het PGO welke gevoeliger zouden zijn voor zelf-gerapporteerde bias dan elektronische vragenlijsten (Jeugdmonitor), al zijn conclusies over dit effect niet eenduidig [37, 38]. Zodoende is er dus niet alleen een verschil in anonimiteit maar ook in de context waarin de vragenlijst wordt afgenomen. De opzet van de huidige studie maakt het niet mogelijk een uitspraak te doen over de precieze rol van deze verschillende mechanismes. Dit neemt echter niet weg dat de resultaten duidelijk laten zien dat de scores uit de verschillende settings niet direct met elkaar kunnen worden vergeleken.

In lijn met eerder onderzoek is de gevonden betrouwbaarheid van de totale probleemscore van de SDQ zowel binnen het PGO als de Jeugdmonitor goed [15]. De betrouwbaarheid van sommige subschalen van de SDQ laat echter te wensen over. Vooral de subschalen gedragsproblemen en in mindere mate pro-sociaal gedrag, problemen met leeftijdgenoten en emotionele symptomen laten een lage betrouwbaarheid zien. Dit geeft aan dat de voorkeur dient te worden gegeven aan de totale probleemscore in plaats van de subschalen wat overeenkomt met het primaire doel van de SDQ als een instrument voor vroegsignalering van psychosociale problematiek. De SDQ dient dus in de eerste plaats vooral om een verhoogd risico op psychosociale problematiek te signaleren en niet zozeer om naar verschillende soorten psychosociale problematiek te differentiëren [15, 23, 39].

Het voorliggende onderzoek kent ook een aantal tekortkomingen. Zo is het voor de beantwoording van de huidige onderzoeksvraag niet mogelijk om een experimenteel design toe te passen, omdat de ware aard van de PGO-setting zich moeilijk laat nabootsen in een experimentele setting. Daarom is er

gebruik gemaakt van een observationeel onderzoek. Hierdoor zijn er echter kleine verschillen tussen de respondenten uit het PGO en de Jeugdmonitor op de achtergrondkenmerken wat mogelijk ook verklaard wordt door het verschil in non-respons. Er is echter eerder aangetoond dat dit verschil in achtergrondkenmerken en non-respons geen invloed had op de resultaten van het huidige onderzoek [40]. Dit onderzoek toonde tevens aan dat de manier van afname, papieren vragenlijsten tegenover een elektronische vragenlijst, geen invloed hadden op de resultaten. Dit gold ook voor het feit dat de Jeugdmonitor klassikaal werd afgenomen. Er is verder een verschil in timing in de twee onderzochte settings. Het lijkt echter niet aannemelijk dat dit een effect heeft gehad op de resultaten, omdat er maar 3 maanden verschil zat in de mediane maand waarin de vragenlijsten zijn afgenomen. Daarnaast moeten de afkappunten niet worden gezien als afkappunten die generaliseerbaar zijn voor het PGO in Nederland aangezien er regionale gegevens zijn gebruikt. Het berekenen van de afkappunten is uitsluitend uitgevoerd om te illustreren wat er gebeurt als er afkappunten van een anonieme setting worden toegepast in een vertrouwelijke setting. Hierbij is het niet van belang of deze afkappunten slechts zijn gebaseerd op percentielscores of dat er ook gebruik is gemaakt van een gouden standaard of klinische populatie [33]. Al verdienen deze twee laatste methodes uiteraard de voorkeur bij de daadwerkelijke bepaling van afkappunten, de toepassing van de verkregen afkappunten binnen de JGZ-setting blijft gekenmerkt door de tekortkomingen zoals zijn beschreven in dit artikel.

Los van bovenstaande resultaten is het van belang te vermelden dat het gebruik van afkappunten altijd ter discussie staat. Hierbij gaat het om het arbitraire karakter van afkappunten en het feit dat psychosociale problematiek niet zo zwart-wit is als kan worden gesuggereerd door een afkappunt, maar in feite een continu spectrum is [32].

Concluderend laat deze studie zien dat, waarschijnlijk door de invloed van sociale wenselijkheid versterkt door de mogelijke persoonlijke consequenties bij een verhoogde score, er verschillen zijn in de scores van de SDQ afgenomen in een anonieme versus vertrouwelijke setting. Deze verschillen zorgen ervoor dat het niet mogelijk is om gemiddelde scores uit beide settings met elkaar te vergelijken dan wel afkappunten een-op-een over te nemen. Voor beleidsdoeleinden zou een vergelijk tussen gegevens van een anonieme setting en vertrouwelijke setting namelijk kunnen zorgen voor een conclusie die niet wordt ingegeven door het werkelijke verschil maar door de setting waarin de gegevens zijn verzameld. Aan de andere kant is het gebruik van anonieme gegevens binnen een vertrouwelijke setting zoals het PGO ook niet direct mogelijk aangezien, de verschillen tussen beide settings resulteren in een andere werking van de afkappunten. Ook al is het gebruik van afkappunten altijd onder-

werp van discussie, als ze worden gebruikt moeten ze werken zoals verwacht [32]. Het is daarom van groot belang dat afkappunten worden bepaald in de setting waarin ze worden toegepast.

Financiering Deze studie werd gefinancierd door de Provincie Limburg (projectnummer CWZW2011/57962), de School for Public Health and Primary Care (CAPHRI; Maastricht), en de GGD Zuid-Limburg.

Literatuur

1. Commissie evaluatie basistakenpakket JGZ. Een stevig fundament: Evaluatie van het basistakenpakket jeugdgezondheidszorg 2013. Utrecht: Commissie evaluatie basistakenpakket JGZ; 2013. pag. 2–13.
2. Reijneveld SA, Vogels AG, Hoekstra F, Crone MR. Use of the pediatric symptom checklist for the detection of psychosocial problems in preventive child healthcare. *BMC Public Health*. 2006;6:197.
3. Tremblay RE. The development of aggressive behaviour during childhood: what have we learned in the past century? *Int J Behav Dev*. 2000;24:129–41.
4. Goodman R. The strengths and difficulties questionnaire: a research note. *J Child Psychol Psychiatry*. 1997;38:581–6.
5. Goedhart A, Treffers F, Widénfelt B van. Vragen naar psychische problemen bij kinderen en adolescenten: de Strengths and Difficulties Questionnaire (SDQ). *MGV*. 2003;58:1018–33.
6. Brugman E, Reijneveld SA, Verhulst FC, Verloove-Vanhorick SP. Identification and management of psychosocial problems by preventive child health care. *Arch Pediatr Adolesc Med*. 2001;155:462–9.
7. Glascoe FP. Early detection of developmental and behavioral problems. *Pediatr Rev*. 2000;21:272–80.
8. Postma S. Richtlijn Vroegsignalering van psychosociale problemen 2008. Bilthoven: RIVM; 2008.
9. Pijpers F. Verantwoord gebruik van vragenlijsten in de jeugdgezondheidszorg. Utrecht: Nederlands Centrum Jeugdgezondheid; 2013.
10. Inspectie voor de Gezondheidszorg. Basisset indicatoren publieke gezondheidszorg 2010. Utrecht: Inspectie voor de Gezondheidszorg; 2010.
11. Engelse O, Zwijgers P. Jeugd In Beeld: Gegevens over de jeugdgezondheidszorg. Utrecht: Nederlands Centrum Jeugdgezondheid; 2013.
12. Vogels AGC, Jacobusse GW, Hoekstra F, Brugman E, Crone M, Reijneveld SA. Identification of children with psychosocial problems differed between preventive child health care professionals. *J Clin Epidemiol*. 2008;61:1144–51.
13. Bell I, Mellor D. Clinical judgements: research and practice. *Aust Psychol*. 2009;44:112–21.
14. Goodman R. The extended version of the strengths and difficulties questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiatry*. 1999;40:791–9.
15. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2001;40:1337–45.
16. Goodman R, Ford T, Meltzer H. Mental health problems of children in the community: 18 month follow up. *BMJ*. 2002;324:1496–7.
17. Theunissen MHC, Wolff M de, Grieken A van, Mieloo C. Handleiding voor het gebruik van de SDQ binnen de Jeugdgezondheidszorg. Vragenlijst voor het signalering

- van psychosociale problemen bij 3- tot 17-jarigen. Leiden: TNO, 2016.
18. Durant LE, Carey MP, Schroder KE. Effects of anonymity, gender, and erotophilia on the quality of data obtained from self-reports of socially sensitive behaviors. *J Behav Med.* 2002;25:438–67.
 19. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant.* 2013;47:2025–47.
 20. He JP, Burstein M, Schmitz A, Merikangas KR. The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. *J Abnorm Child Psychol.* 2013;41:583–95.
 21. Niclasen J, Teasdale TW, Andersen AM, Skovgaard AM, Elberling H, Obel C. Psychometric properties of the Danish strength and difficulties questionnaire: the SDQ assessed for more than 70,000 raters in four different cohorts. *PLoS ONE.* 2012;7:e32025.
 22. Mieloo C, Raat H, Oort F van, et al. Validity and reliability of the strengths and difficulties questionnaire in 5–6 year olds: differences by gender or by parental education? *PLoS ONE.* 2012;7:e36805.
 23. Stone LL, Otten R, Engels RC, Vermulst AA, Janssens JM. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clin Child Fam Psychol Rev.* 2010;13:254–74.
 24. Sanne B, Torsheim T, Heiervang E, Stormark KM. The strengths and difficulties questionnaire in the Bergen child study: a conceptually and methodically motivated structural analysis. *Psychol Assess.* 2009;21:352–64.
 25. Roy B van, Veenstra M, Clench-Aas J. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. *J Child Psychol Psychiatry.* 2008;49:1304–12.
 26. Lamberts PHG, Steenbakkens M, Thijssen NHB, et al. Een gezonde kijk op Zuid-Limburg. Regionale Volksgezondheid Toekomst Verkenning 2010. Maastricht: GGD Zuid Limburg; 2011.
 27. CBS. Key figures for postcode areas. Heerlen: CBS; 2009.
 28. Knol F. Neighbourhood status development in the Netherlands. Den Haag: SCP; 2012.
 29. CBS. Foreign background. Den Haag: CBS; 2014.
 30. Sijtsma K. Over misverstanden rond Cronbachs alfa en de wenselijkheid van alternatieven. *Psycholoog.* 2009;44:561–7.
 31. Field A. *Discovering statistics using IBM SPSS statistics.* London: SAGE; 2013.
 32. Miles J, Shevlin M. *Applying regression and correlation: a guide for students and researchers.* London: SAGE; 2001.
 33. Goodman A, Goodman R. Strengths and difficulties questionnaire as a dimensional measure of child mental health. *J Am Acad Child Adolesc Psychiatry.* 2009;48:400–3.
 34. Vogels AGC, Siebelink BM, Theunissen MC, Wolff MS de, Reijneveld SA. Vergelijking van de KIVPA en de SDQ als signaleringsinstrument voor problemen bij adolescenten in de Jeugdgezondheidszorg. Leiden: TNO; 2011.
 35. Looij-Jansen PM van de, Goldschmeding JEJ, Wilde EJ de. Comparison of anonymous versus confidential survey procedures: effects on health indicators in Dutch adolescents. *J Youth Adolesc.* 2006;35:652–8.
 36. Aquilino WS, Wright DL, Supple AJ. Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use. *Subst Use Misuse.* 2000;35:845–67.
 37. Griesler PC, Kandel DB, Schaffran C, Hu M-C, Davies M. Adolescents' inconsistency in self-reported smoking: a comparison of reports in school and in household settings. *Public Opin Q.* 2008;72:260–90.
 38. Looij-Jansen PM van de, Wilde EJ de. Comparison of web-based versus paper-and-pencil self-administered questionnaire: effects on health indicators in Dutch adolescents. *Health Serv Res.* 2008;43:1708–21.
 39. Widenfelt BM van, Goedhart AW, Treffers PD, Goodman R. Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *Eur Child Adolesc Psychiatry.* 2003;12:281–9.
 40. Hoofs H, Jansen NW, Mohren DC, Jansen MW, Kant IJ. The Context Dependency of the Self-Report Version of the Strength and Difficulties Questionnaire (SDQ): A Cross-Sectional Study between Two Administration Settings. *PLoS ONE.* 2015;10:e120930.

H. Hoofs, Promovendus

N. Jansen, Hoofddocent

D. Mohren, Onderzoeker

M. Jansen, Hoogleraar

IJ. Kant, Hoogleraar