

BIJLAGE 3: POWERBEREKENINGEN BLOOTSTELLING-EFFECT ASSOCIATIES

Om een inschatting te kunnen maken van het benodigde aantal deelnemers in de onderzoeksgroep voor het uitvoeren van de blootstelling-effect associaties werden powerberekeningen uitgevoerd. Deze werden apart uitgevoerd voor binaire effectmerkers (ja/nee uitkomst zoals bijvoorbeeld het hebben van astma of een allergie) en voor continue effectmerkers (bijv. cholesterolgehalte, hormoonconcentraties).

1. Binaire effectmerkers

Voor de **binaire merkers** (later geanalyseerd met een logistisch regressie model) is als input voor de berekeningen nodig: de odds ratio (OR) die samengaat met een bepaalde toename in de blootstellingsmerkers en de prevalentie van de effectmerker.

- Vaak wordt de OR voor een toename in blootstellingmerker van P25 naar P75 gebruikt. De power berekening (in SAS) gebruikt geen input over het effect van andere verklarende variabelen (bijv leeftijd, geslacht, ...) etc. Er is wel de onderliggende veronderstelling dat de blootstellingsmerker onafhankelijk is van deze verklarende variabelen. Dat is niet noodzakelijk het geval. Hoe sterker de correlatie tussen de blootstellingmerker en de andere covariaten (R^2 variantie in blootstelling verklaard door de andere covariaten) hoe groter de steekproef moet zijn om eenzelfde power te bekomen: $n' = n/(1 - R^2)$. Onderstaande tabel toont hoeveel extra observaties nodig zijn om eenzelfde power te bekomen. Als bijv. 5% van de variantie in blootstellingmerker verklaard wordt door de andere verklarende variabelen; dan zijn 5.26% meer deelnemers nodig om eenzelfde power te hebben dan wanneer de blootstellingmerker en de verklarende variabelen onafhankelijk zijn.

R^2	% extra observaties
0,05 (5 procent)	5,26
0,1	11,11
0,15	17,65
0,2	25,00
0,5	100,00

- De prevalentie wanneer alle verklarende variabelen gelijk zijn aan de gemiddelde waarde. Hiervoor kunnen we het % deelnemers met voorkomen van de effectmerker nemen.

Onderstaande figuren tonen dan hoe de power verandert als functie van de prevalentie en de OR. Voor de prevalenties zijn er lijnen voor een voorkomen van 10%, 20%, ...50%. De lijn voor een prevalentie van 90% valt samen met deze van 10%; die van 80% met deze van 20% etc.

De OR werd gevarieerd van 0.1 tot 5. Veel hogere OR's lijken onwaarschijnlijk. Een OR=1 geeft dus aan dat de blootstelling niet samenhangt met het effect. De eerste figuur geeft de power weer voor een steekproef van 150 deelnemers; de tweede voor een steekproef van 300 deelnemers.

Met 300 deelnemers kunnen OR's < 0.7 of > 1.6-1.7 gedetecteerd worden (tenzij de prevalentie lager is dan 20%). Met 150 deelnemers zijn dit OR < 0.5 of > 2 en met 75 deelnemers zijn dit OR's < 0.3 of > 2.7-3.2.

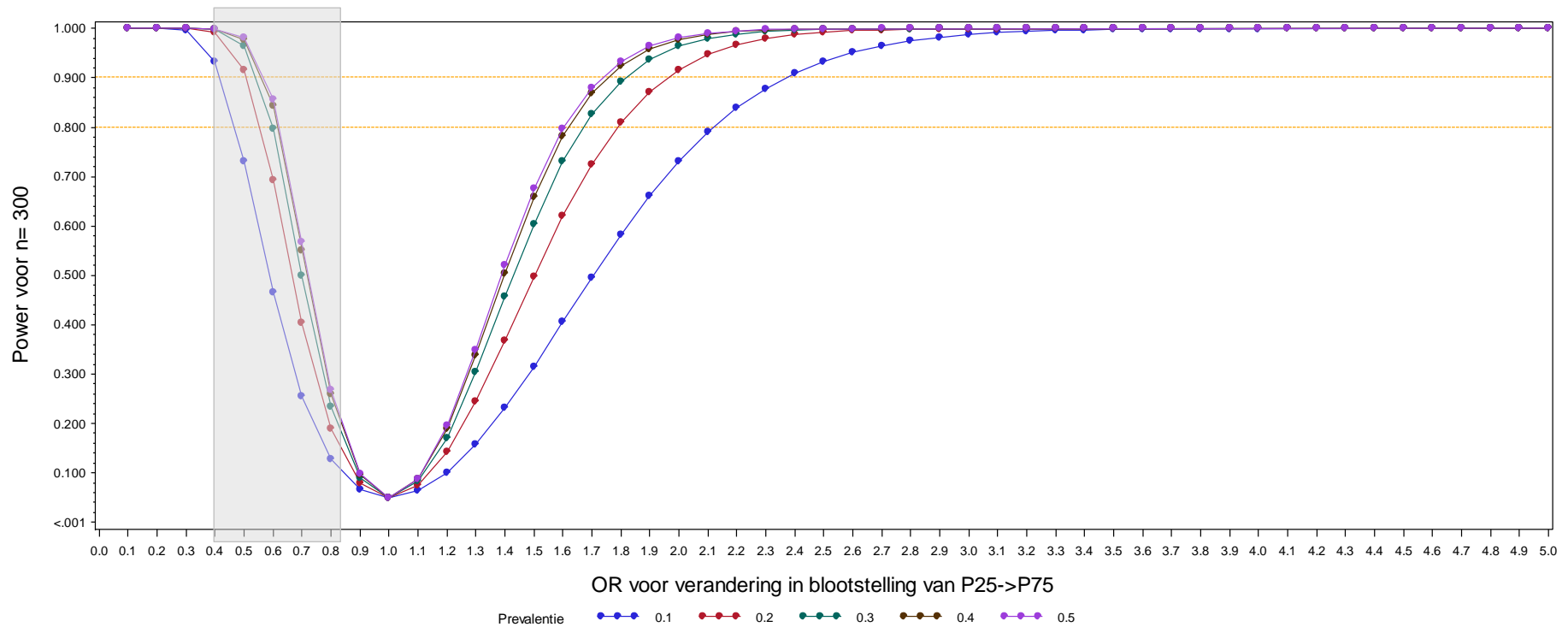
Een aantal van de OR's die in de rapporten van het Steunpunt Milieu en Gezondheid gerapporteerd worden, zitten in deze ranges. Voorwaarde is nu natuurlijk wel nog dat de toename van P25-P75 die in die verwerkingen gezien is; ook in de nieuwe resultaten data zit.

Studie	IQR	prevalentie	effect
FLEHS II newborns	PFOS: 2 PFOA: 0.9	Astma ooit: 25% Astma huidig: 17% Astma dokter: 7% Allergie producten: 30%	PFOS-astma ooit: OR=0.62 PFOA-astma ooit: OR=0.64 PFOA-astma huidig: OR=0.61 PFOS-allergie producten: OR=0.51
FLEHS II Menen	PFOS: 3.8 PFOA: 0.87	Astma ooit: 17% Astma huidig: 12.5% Astma dokter: 9% Allergie voeding/medicatie: 21% Allergie producten: 16%	PFOS-allergie voeding/medicatie: OR=0.39 PFOS-allergie producten: OR=0.53
FLEHS II adults (20-40j)	PFOS: 7.9 PFOA: 2	Astma ooit: 20% Astma huidig: 15% Astma dokter: 7%	
FLEHS III newborns	PFOS: 0.95 PFOA: 0.68 PFHxS: 0.28 PFNA: 0.17	t-t-p huidig: 10% t-t-p ooit: 13% miskraam: 26% huidig astma: 16% astma ooit: 24% hooikoorts: 23% allergie producten: 16%	PFOS-t-t-p: OR=0.79 PFNA-t-t-p: OR=0.83 PFOS-t-t-p ooit: OR=0.40 PFOS-miskraam: OR=0.68 PFOA-miskraam: OR=0.63 PFOS-huidig astma: OR=0.55 PFOS-astma ooit: OR= 0.62 PFHxS-astma ooit: OR=0.66 PFOS-hooikoorts: OR=0.51 PFHxS-hooikoorts: OR=0.55 PFHxS-allergie producten: OR=0.52
FLEHS III adults (50-65j)	PFOS: 5.93 PFOA: 1.56 PFHxS: 1.34 PFNA: 0.58	Huidig astma: 15%	PFOS-huidig astma: OR=0.56 PFOA-huidig astma: OR=0.48
FLEHS IV adolescents	PFOS: 1.7 PFOA: 0.46 PFHxS: 0.3 PFNA: 0.2		
3M fase 1 volledige groep (12-90j, n=796)	PFOS _{lin} : 12.61 PFOS _{l+v} : 33.3 PFOA _{lin} : 1.1 PFOA _{l+v} : 1.1 PFHxS _{lin} : 1.87 PFHxS _{l+v} : 1.89 PFNA: 0.35		
3M fase 1 volledige groep (12-19j, n=34)	PFOS _{lin} : 39.11 PFOS _{l+v} : 37.38 PFOA _{lin} : 0.59 PFOA _{l+v} : 0.61 PFHxS _{lin} : 1.06 PFHxS _{l+v} : 1.04 PFNA: 0.17		

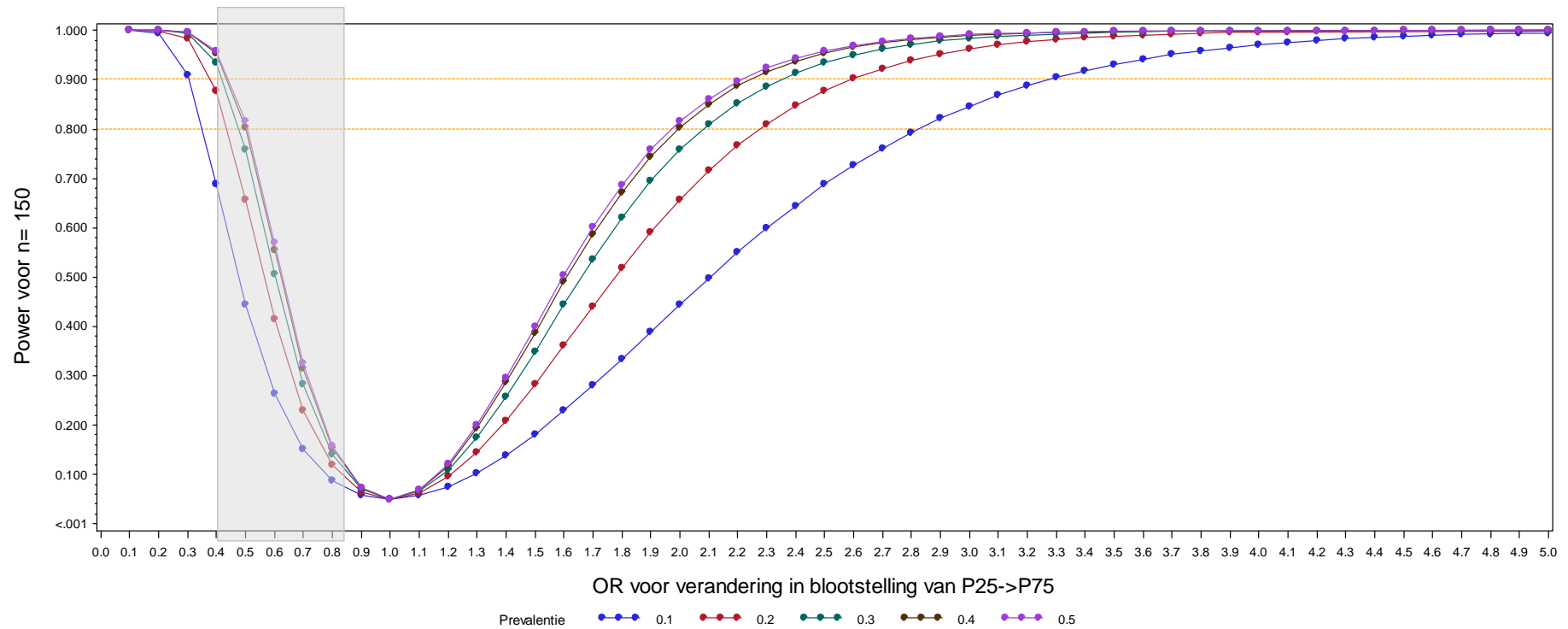
3M fase 1 volledige groep (20-49j, n=348)	PFOS _{lin} : 7.16 PFOS _{I+v} : 22.02 PFOA _{lin} : 0.80 PFOA _{I+v} : 0.82 PFHxS _{lin} : 1.07 PFHxS _{I+v} : 1.08 PFNA: 0.22		
---	---	--	--

De range OR's die in de Steunpunt-rapporten werden gerapporteerd, zijn weergegeven als een grijs vlak in de figuren 1-3. Uit deze figuren blijkt dat bij een studiepopulatie van 300 de effecten met OR tussen 0.4 en 0.7 kunnen worden opgepikt, bij 150 deelnemers zijn dit de effecten met een OR tussen 0.4 en 0.6, maar bij 75 deelnemers is het weinig waarschijnlijk dat de OR's die in de Steunpunt-studies werden gerapporteerd zouden worden opgepikt.

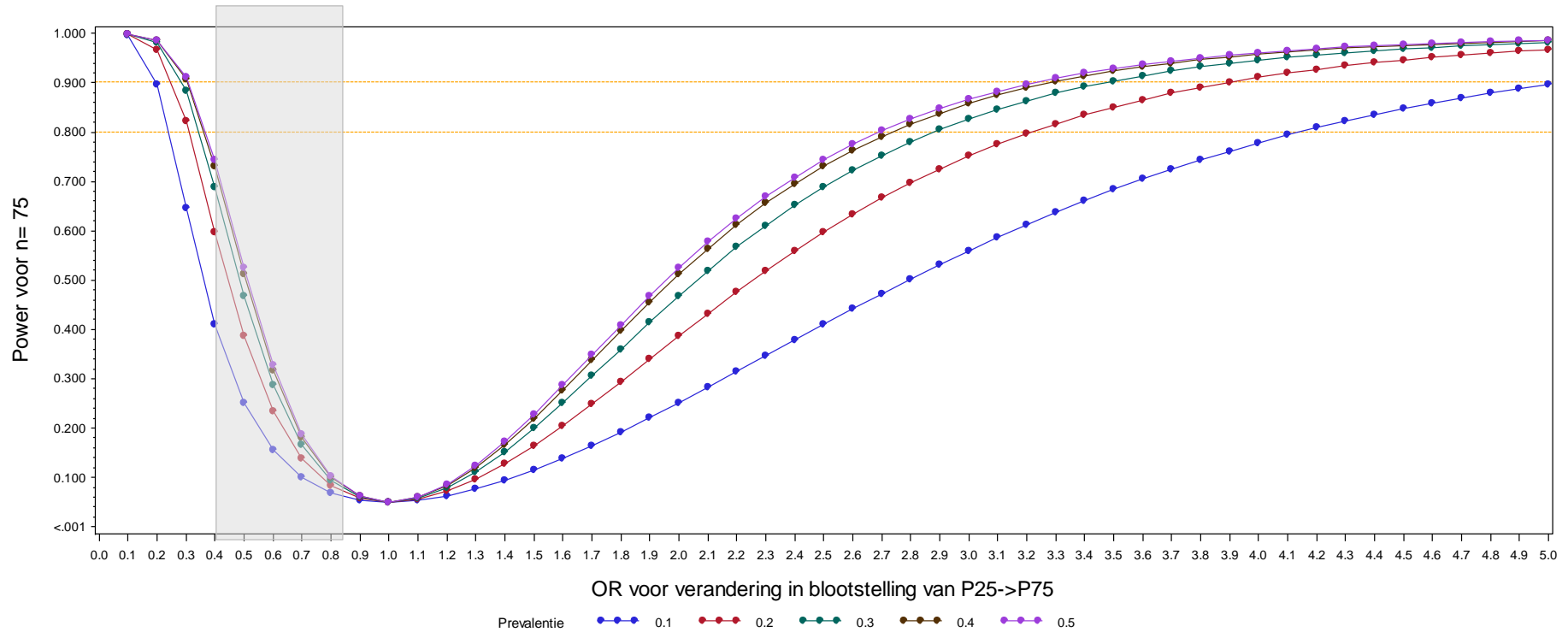
Figuur 1: Powerberekeningen bij n=300 deelnemers in functie van variatie in odds ratio (OR) voor binaire effectmerkers bij een verandering in blootstelling van de P25 naar de P75 voor 10%, 20%, 30%, 40% en 50% prevalentie van het effect. Het grijze vlak toont de OR's waargenomen in de gerapporteerd associaties met PFAS in de resultaten van het Steunpunt Milieu en Gezondheid.



Figuur 2: Powerberekeningen bij n=150 deelnemers in functie van variatie in odds ratio (OR) voor binaire effectmerkers bij een verandering in blootstelling van de P25 naar de P75 voor 10%, 20%, 30%, 40% en 50% prevalentie van het effect. Het grijze vlak toont de OR's waargenomen in de gerapporteerd associaties met PFAS in de resultaten van het Steunpunt Milieu en Gezondheid.



Figuur 3: Powerberekeningen bij n=75 deelnemers in functie van variatie in odds ratio (OR) voor binaire effectmerkers bij een verandering in blootstelling van de P25 naar de P75 voor 10%, 20%, 30%, 40% en 50% prevalentie van het effect. Het grijze vlak toont de OR's waargenomen in de gerapporteerd associaties met PFAS in de resultaten van het Steunpunt Milieu en Gezondheid.

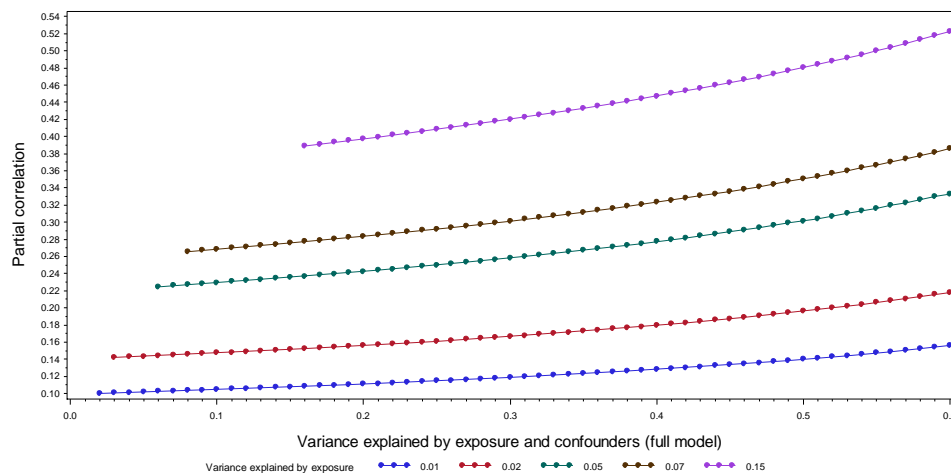


2. Continue effectmerkers

Voor de **continue blootstellingsmerkers** kan gewerkt worden ofwel met de toename in verklaarde variabiliteit door de blootstellingsmerker toe te voegen aan een model met al een set van covariaten, ofwel met de partiële multiple correlatie/determination coëfficiënt (die hangen samen). Het voordeel is dat er hier geen veronderstellingen moeten gemaakt worden over de vorm van de associatie (lineair kwadratisch); continue of categorische verklarende variabele, interactie, ln transformatie, ...

- Stel dat een model gefit wordt $y \sim x$, waarbij x de set confounders is zoals leeftijd, geslacht, Dit model heeft een $R^2 = R^2_{y|x}$
- Het model $y \sim x + z$, met z de blootstellingsmerker heeft een $R^2 = R^2_{y|xz}$
- De partiële multiple correlatie R is dan gelijk aan $\sqrt{((R^2_{y|xz} - R^2_{y|x}) / (1 - R^2_{y|x}))}$. Deze correlatie is dus de correlatie tussen de effect merker (y) en de blootstelling (z) na correcties voor de andere covariaten (x).

Deze maten zijn niet opgenomen in de resultaten gerapporteerd in de Steunpunt-rapporten of in de doorgenomen literatuur. Dus het is moeilijk om voeling te hebben met wat kan verwacht worden. Onderstaande figuur toont de relatie tussen de partiële correlatiecoëfficiënt en de verklaarde variantie van het full model (met x en z) en wat de blootstellingsmerker daarvan voor zijn rekening neemt.



Op basis van deze maten geeft Cohen (1988)¹ een effectsize $f^2 = R^2/(1+R^2)$ waarbij een $f^2 < 0.02$ 'very weak' association is; $0.02 < f^2 < 0.15$ 'weak'; $0.15 < f^2 < 0.35$ 'medium'; en $f^2 > 0.35$ 'substantial'. Onderstaande tabel geeft weer welke partial coefficient of determination (R^2), partial correlation coefficient (R) met deze effectsizes (f^2) overeenkomen.

f^2	R^2	R
0,02	0,019608	0,140028
0,15	0,130435	0,361158
0,35	0,259259	0,509175

¹ [Statistical Power Analysis for the Behavioral Sciences | Jacob Cohen | \(taylorfrancis.com\)](https://www.taylorfrancis.com/books/9780429015311/statistical-power-analysis-behavioral-sciences/jacob-cohen/)

Om de power te berekenen heeft ook het aantal vrijheidsgraden een effect (m.a.w. het aantal verklarende variabelen in het model), maar dat effect is vrij klein hier.

Ook hier werd gevarieerd met de effect size/partiële correlatie en werd de power bekeken voor een steekproef van 75, 150 of 300 deelnemers. Met 150 deelnemers kunnen partiële correlaties vanaf 0.23/0.26 gedetecteerd worden met een power van 80/90%. Dit zijn volgens de indeling van Cohen medium effect sizes. In dat geval verklaart de blootstellingsmerker 5-7% van de variantie. Met 300 deelnemers kunnen partiële correlaties vanaf 0.16/0.19 met een power van 80/90% gedetecteerd worden. Dit gaat al in de richting van een 'weak' associatie, dus ook de minder sterke associaties kunnen worden opgepikt. In dit geval verklaart de blootstellingmerker 2.5-4% van de variantie. Indien de steekproef bestaat uit 75 deelnemers (bijvoorbeeld enkel vrouwen of mannen van een groep van 150 deelnemers) dan kunnen partiële correlaties vanaf 0.31/0.36 gedetecteerd worden met een power van 80/90%, wat overeen komt met 'substantial' effect sizes. Dan kunnen vooral de sterkere associaties gedetecteerd worden. De blootstellingsmerker verklaart dan ongeveer 10-15% van de variantie.

Omdat we geen inzicht hebben op realistische partiële correlatie coëfficiënten voor de Vlaamse HBM-data werden drie voorbeelden opnieuw berekend met beschikbare gegevens uit de Steunpunt-dataset.

STP4: data van ± 606 jongeren over heel Vlaanderen

- CHOL (cholesterol) en PFOS/PFOA

STP2: data van ± 200 jongeren uit de regio Mene

- CHOL (cholesterol) en PFOS/PFOA
- TSH (thyroid stimulating hormone) en PFOS/PFOA
- FRE2 (vrij oestradiol) en PFOS/PFOA voor de jongens

PFOS, PFOA, CHOL en TSH werden volgens de natuurlijke logaritme getransformeerd. FRE2 wordt niet getransformeerd.

Als verklarende variabelen werden volgende gebruikt (gebaseerd op variabelen gebruikt in de Steunpuntrapporten en in de Veneto-studie²):

- CHOL:
 - STP4: geslacht, fysieke activiteit, roken, herkomst op basis van geboorteland, consumptie van fruit, groenten, chocolade, eieren, vis bij de warme maaltijd, brood, graanproducten, alcohol, fastfood tijdens de voorbije 3 dagen, gefrituurd voedsel tijdens de voorbije 3 dagen, hoogste opleiding in het gezin, leeftijd en body mass index
 - STP2: leeftijd, geslacht, ziek tijdens de voorbije 14 dagen, body mass index
- TSH: leeftijd, geslacht, ziek tijdens de voorbije 14 dagen, body mass index, fysieke activiteit, roken, herkomst op basis van geboorteland, consumptie van groenten, fruit, melk, kaas, vlees, eieren, vis, alcohol, hoogste opleiding in het gezin

² [Associations between perfluoroalkyl substances and lipid profile in a highly exposed young adult population in the Veneto Region - ScienceDirect](#)
[Perfluoroalkyl substances and thyroid stimulating hormone levels in a highly exposed population in the Veneto Region - ScienceDirect](#)

- FRE2: leeftijd, tijdstip bloedafname, roken, body mass index, seizoen bloedafname, ziek tijdens de voorbije 14 dagen

Met een beperktere of uitgebreidere set werken heeft vooral effect op de totaal verklaarde variabiliteit, maar niet zo heel veel op de partiële correlatiecoëfficiënt.

De blootstellingsmerkers, op ln schaal, werden als kwadratische termen in het model opgenomen.

Dan werd volgende info i.v.m. R^2 waarden en partiële correlatiecoëfficiënten bekomen (Tabel 1). Deze in vet/scheef gedrukt zijn ook significante associaties in deze ruwe verwerking.

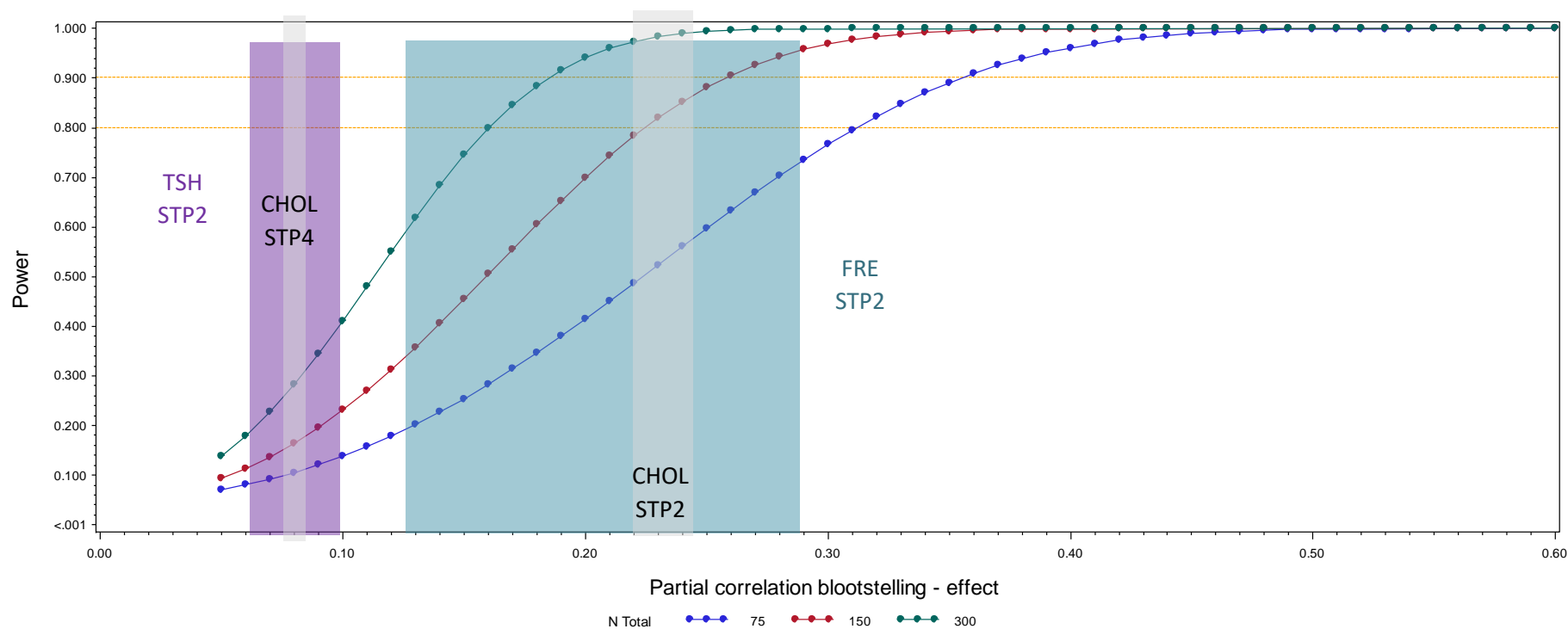
Interpretaties (Tabel 1):

- TSH: In de resultaten van STP2 (200 deelnemers) kan geen significante associatie tussen TSH en PFOA of PFOS worden aangetoond. Om partiële correlaties van 0.05-0.10 met een power van 80% te kunnen detecteren, zouden 780-3135 deelnemers nodig zijn.
- CHOL: Voor de data van STP 2 (200 deelnemers) kunnen significante associaties met CHOL worden gedetecteerd met PFOA en met PFOS. Op basis van de data van STP 4 (600 deelnemers) is er een ander beeld. De partiële correlaties voor STP4 liggen veel lager dan deze op basis van de data van STP2. De medianen en IQR (interkwartielafstand) voor CHOL en voor PFOA en PFOS zijn in STP2 hoger dan deze in STP4.
- FRE2: In de resultaten van STP2 (100 deelnemers, enkel de jongens) kan een significante associatie worden vastgesteld tussen vrij oestradiol en PFOA, maar niet met PFOS. De partiële correlatie voor PFOS ligt lager dan voor PFOA.

Besluit:

Met 300 deelnemers kunnen partiële correlaties vanaf 0.16/0.19 met een power van 80/90% gedetecteerd worden. Dit wil zeggen dat ook de minder sterke associaties kunnen worden opgepikt. In dit geval verklaart de blootstellingmerker 2.5-4% van de variantie. Voor analyses die voor jongens en meisjes apart moeten gedaan worden, kunnen met 150 deelnemers partiële correlaties vanaf 0.23/0.26 gedetecteerd worden met een power van 80/90%. Dan kunnen dus associaties van medium sterkte worden opgepikt. In dat geval verklaart de blootstellingsmerker 5-7% van de variantie. De vier voorbeelden op basis van bestaande gegevens in het Steunpunt Milieu en Gezondheid tonen aan dat deze partiële correlaties realistisch zijn en maar ook afhankelijk zijn van de spreiding van zowel de blootstelling als de effectmerkers in de deelnemersgroep.

Figuur 4: Powerberekeningen in functie van partiële correlatie voor continue effectmerkers bij een verandering in blootstelling van de P25 naar de P75 voor populatiegroottes van 75, 150 en 300 deelnemers. *Paarse balk*: bereik van de partiële correlatiecoëfficiënten voor TSH met PFOA en PFOS in het 2^{de} Steunpunt Milieu en Gezondheid, *blauwe balk*: bereik van de partiële correlatiecoëfficiënten voor vrij oestradiol (FRE2) met PFOA en PFOS in het 2^{de} Steunpunt Milieu en Gezondheid, *grijze balken*: bereik van de partiële correlatiecoëfficiënten voor cholesterol met PFOA en PFOS in het 2^{de} en in het 4^{de} Steunpunt Milieu en Gezondheid,



Tabel 1: R^2 en partiële correlatiecoëfficiënten voor de associaties tussen cholesterol (CHOL_In), thyroid stimulerend hormoon (TSH_In) en vrij oestradiol (FRE) met PFOA en PFOS in serum.

Effect merker	STP	R^2 model met enkel confounders/covariaten	+ PFOA (ln kwadratisch)	→ Partiele correlatiecoëfficiënt	+ PFOS (ln kwadratisch)	→ Partiele correlatiecoëfficiënt
CHOL_In	2	0.141	0.189	0.219	0.180	0.244
	4	0.106	0.112	0.074	0.110	0.083
TSH_In	2	0.180	0.183	0.102	0.188	0.060
FRE	2	0.367	0.416	0.288	0.376	0.125

Studie in opdracht van:

Studie in opdracht van: