

SYNTHESE

MULTICRITERIA BESLISSINGSANALYSE BIJ DE SELECTIE VAN PRIORITAIRE AANDOENINGEN BINNEN DE NEONATALE BLOEDSCREENING



SYNTHESE

MULTICRITERIA BESLISSINGSANALYSE BIJ DE SELECTIE VAN PRIORITAIRE AANDOENINGEN BINNEN DE NEONATALE BLOEDSCREENING

CHRIS DE LAET, GERMAINE HANQUET, ERIK HENDRICKX



■ VOORWOORD

Je zou anno 2016 kunnen verwachten dat de gezondheidszorg in Europa zo stilaan een zekere graad van harmonisering heeft bereikt – we hebben in zowat elke discipline ‘*European societies*’ die ijverig richtlijnen en consensusdocumenten produceren, die ze toch allemaal baseren op dezelfde internationale bronnen van evidence-based informatie, niet? Welnu, wat betreft de screening van pasgeborenen op aangeboren metabole ziekten blijkt niets minder waar. Al naargelang het land waar je als kers-vers Europeaanje het levenslicht ziet, wordt je hielprik op heel andere aandoeningen getest. Qua aantal geteste ziekten is de range enorm, gaande van één of twee tot een dertigtal. Akkoord, niet alle Europeaanjes komen uit een vergelijkbare genen-pool, en het heeft niet zoveel zin om op thalassemie te testen in Lapland, ... al kunnen bepaalde migrantenpopulaties dit verhaal behoorlijk complexer maken. Maar ook tussen landen – of regio’s – met vergelijkbare aantallen geteste ziekten, kunnen de panels erg van mekaar verschillen.

Het kiezen van een ‘optimaal’ screeningspanel is nu eenmaal een complexe kwestie, waarbij een heel aantal verschillende dimensies tegelijkertijd moeten worden afgewogen. En we kennen ze allemaal, de commissievergaderingen waarin dergelijke complexe knopen moeten worden doorgehakt: hoe de discussie voortdurend van de ene dimensie naar de andere dreigt te springen, hoe het debat gedomineerd kan worden door ‘opinion leaders’, hoe de verschillende stakeholders elk hun eigen belichting van het probleem naar voor schuiven en tenslotte niemand het overzicht kan bewaren. En dus eindigt dit onvermijdelijk met in elk land of regio een ander lijstje – daarom niet totaal ongefundeerd – maar alleszins erg contingent en weinig consistent.

Her en der groeit het besef dat de besluitvorming rond deze screenings hygiënischer zou kunnen verlopen – en wij vonden het voorbeeld van Québec wat dat betreft erg inspirerend. In de studie die voor u ligt, hebben we dan ook geprobeerd om hun model van multiple-criteria decision analysis (MCDA) te testen op zijn bruikbaarheid voor onze Belgische (lees Vlaamse, Waalse en Brusselse) realiteit. Dit is de tweede oefening van haar soort in het Belgische gezondheidssysteem, want daarnaast zijn we samen met het RIZIV ook met een MCDA model aan de slag in het dossier van de *unmet medical needs*.

We mogen hopen dat dit de prille tekenen zijn van een belangrijke cultuuromslag, waarbij men het aandurft om – met de ondersteuning van dergelijke tools – de besluitvorming in zulke complexe kwesties objectiever en, vooral, veel transparanter te maken.

Christian LÉONARD
Adjunct Algemeen Directeur

Raf MERTENS
Algemeen Directeur



■ **SYNTHESE**

INHOUDSTAFEL

■	VOORWOORD.....	1
■	SYNTHESE	2
1.	NEONATALE BLOEDSCREENINGSPROGRAMMA'S (NBS).....	3
2.	SCOPE EN METHODOLOGIE	5
3.	SELECTIE EN DEFINITIE VAN DE CRITERIA.....	6
4.	BEPALING VAN HET GEWICHT VAN ELK CRITERIUM	8
5.	VERZAMELEN VAN INFORMATIE OVER ELKE AANDOENING	12
6.	HET TOEKENNEN VAN DE SCORES PER AANDOENING.....	12
7.	Globale Eindscore per ziekte	16
8.	CONCLUSIE	18
9.	KERNBOODSCHAPPEN EN LESSEN VOOR DE TOEKOMST	18
■	AANBEVELINGEN	20



1. NEONATALE BLOEDSCREENINGSPROGRAMMA'S (NBS)

Het neonatale bloedscreeningsprogramma (NBS) is een volksgezondheidsprogramma dat alle zuigelingen kort na de geboorte systematisch screent op een reeks van behandelbare aandoeningen die in de neonatale periode klinisch niet meteen zichtbaar zijn. De meeste aandoeningen in NBS programma's zijn aangeboren stofwisselingsziekten waarvan de eerste symptomen verschijnen in de neonatale periode of de vroege kindertijd, terwijl er dan al onherstelbare schade kan zijn aangericht. Door deze aandoeningen op te sporen voordat symptomen opduiken zijn een vroege diagnose en tijdige interventie mogelijk. Neonatale screening wordt ook vaak neonatale hielprik of Guthrietest genoemd. Daarbij wordt er een klein bloedstaal van de baby genomen, en dit bloed wordt aangebracht op een voorgedrukt kaartje (bekend als de Guthriekaart).

Internationaal is er geen duidelijke consensus over de op te sporen aangeboren ziekten. Daardoor verschilt hun aantal (van geen tot meer dan veertig) en aard erg van land tot land.

In België ging het NBS-programma in 1968 van start, met een systematische screening van alle pasgeborenen op fenyketonurie. Zes andere aandoeningen werden geleidelijk toegevoegd aan het programma.

In de vroege jaren tachtig werd de verantwoordelijkheid van het programma overgedragen naar de Vlaamse Gemeenschap (VG) en de Communauté française (later kreeg deze de naam Fédération Wallonie-Bruxelles (FWB)).^a Sindsdien beslist elke gemeenschap zelf op welke aandoeningen wordt gescreend. Ze baseren zich daarvoor op hun juridische criteria en de

adviezen van hun adviescommissies. Het gevolg is dat de twee gemeenschappen niet dezelfde aandoeningen opsporen. Vandaag wordt er in de VG op 11 en in de FWB op 13 aandoeningen gescreend. Negen worden er in beide gemeenschappen opgespoord (Tabel 1).

Beide gemeenschappen gebruiken gelijkaardige technologieën voor de screening, maar de testen worden in afzonderlijke referentiecentra uitgevoerd (twee in VG en drie in FWB). De meeste aangeboren stofwisselingsziekten (maar niet allemaal) worden tegenwoordig opgespoord door de tandem massaspectrometrie (MS/MS). Bij voldoende volume is dit een relatief goedkope labotechniek waar door de analyse van een enkel bloedstaal een groot aantal stofwisselingsziekten gelijktijdig en snel kunnen worden opgespoord. Deze techniek meet concentraties van een groot aantal metabolieten en, als de test positief is, moet ze worden gevolgd door bevestigende testen voor de einddiagnose. In de VG bedraagt de totale kostprijs voor de primaire screening (inclusief kaartjes, verzenden, meedelen van de resultaten en promotiemateriaal) ongeveer 20 € per baby.

De MS/MS-techniek kan potentieel een groot aantal aandoeningen opsporen. Daarom moet men beslissen welke aandoeningen opgenomen moeten worden in de screeningsprogramma's. Daarbij moet men een aanvaardbaar evenwicht tussen voordelen en risico's nastreven. Mogelijke voordelen zijn dat vroegtijdige opsporing en een effectieve interventie de aandoening, sequelae en soms zelfs vroegtijdig overlijden kunnen voorkomen. De belangrijkste risico's zijn de gevolgen van vals negatieve en vals positieve resultaten, met respectievelijk een foute geruststelling of onnodige bezorgdheid en kosten tot gevolg.

Het doel van deze KCE-studie is de gemeenschappen te helpen bij de besluitvorming over de aandoeningen die opgenomen kunnen worden in het NBS-programma.

^a In FWB, wordt het NBS programma sinds 2015 beheert door het Office de la Naissance et de L'Enfance (ONE)



Tabel 1– Ziekten opgenomen in neonatale bloedscreeningsprogramma's op gemeenschapsniveau (februari 2016)

Ziekte	Afkorting	In de Vlaamse Gemeenschap	In de Fédération Wallonie Bruxelles	Opgenomen in deze pilootstudie
<i>Stofwisselingsaandoeningen</i>				
Biotinidase deficiëntie	LMCD	Ja	Nee	Ja
Galactosemie	GAL	Nee	Ja	Ja
Glutaaracidemie type 1	GA I	Ja	Ja	
Homocystinurie	HCY	Nee	Ja	Ja
Isovaleriaanacidemie	IVA	Ja	Ja	
Maple syrup urine ziekte	MSUD	Ja	Ja	
Medium chain acyl-CoA dehydrogenase deficiëntie	MCAD	Ja	Ja	
Methylmalonacidemie	MMA	Ja	Ja	
Multipel acyl-CoA dehydrogenase deficiëntie	MADD	Ja	Ja	
Fenylketonurie	PKU	Ja	Ja	
Propionacidemie	PA	Ja	Ja	
Tyrosinemie Type I	TYR I	Nee	Ja	Ja
Very long chain CoA dehydrogenase deficiëntie	VLCAD	Nee	Ja	Ja
<i>Endocriene stoornissen</i>				
Congenitale hypothyroidie	CHT	Ja	Ja	
Congenitale bijnierhyperplasie	CAH	Ja	Nee	Ja



2. SCOPE EN METHODOLOGIE

In deze studie doen we een piloottest van een Multi-Criteria Decision Analysis (MCDA) methode gericht op ziekteprioritering. Deze methode werd in 2013 toegepast in Québec door het Institut National d'Excellence en Santé et en Services Sociaux (INESSS).

In een MCDA selecteert men a priori een aantal criteria die elk een specifiek gewicht krijgen. De aandoeningen waarop kan worden gescreend krijgen vervolgens een score voor elk van deze criteria en worden gerangschikt op basis van een globale eindscore.

In onze pilootstudie bestaat dit uit de volgende stappen:

1. Selecteren en nader definiëren van criteria voor de evaluatie van aandoeningen.
2. Definiëren van het gewicht van elk criterium, dat verder wordt gebruikt bij de berekening van een globale eindscore per aandoening. Het gewicht voor een criterium wordt uitgedrukt aan de hand van een LIKERT-schaal van 1 tot 4 (Tabel 2).
3. Voorbereiden van wetenschappelijke informatie voor elke aandoening en voor elk criterium in een korte samenvattende ziektefiche.
4. Per aandoening de score bepalen voor elk criterium. Ook deze score wordt uitgedrukt aan de hand van een LIKERT-schaal van 1 tot 4 (Tabel 2).
5. Berekenen van een globale eindscore per aandoening: in het algemeen worden hiervoor de gewichten van het criterium vermenigvuldigd met de score per ziekte. Deze producten (gewogen scores) worden opgeteld per ziekte. Afhankelijk van de methode geeft dit ofwel direct een globale eindscore per ziekte, ofwel een eindscore voor elk van de beoordeelaars. In dit laatste geval worden de individuele eindscores samengevat om een globale eindscore per ziekte te berekenen wat moet toelaten deze aandoeningen te rangschikken.^b

Tabel 2 – Beschrijving van de LIKERT-schaal om het criteriumgewicht en de aandoening te scoren

Schaal	Gewicht voor het criterium	Score voor de aandoening voor dit criterium
1	niet relevant	zeer laag
2	weinig relevant	eerder laag
3	relevant	eerder hoog
4	zeer relevant	zeer hoog
0		niet te beantwoorden, ontbrekende gegevens

Stap 1 werd uitgevoerd door het KCE-team samen met een stuurgroep van beleidsverantwoordelijken uit beide gemeenschappen, experts uit de NBS-labo's en klinische experts, patiënten vertegenwoordigers en ethici (in totaal 11 beoordeelaars). Stappen 2 en 4 werden enkel door de leden van de stuurgroep uitgevoerd, terwijl stappen 3 en 5 werden uitgevoerd door het KCE-team en later besproken met de stuurgroep.

De stuurgroep en het KCE-team beslisten om voor deze pilootstudie de zes aandoeningen te selecteren waarop enkel in één van beide gemeenschappen wordt gescreend (). Omdat de scope van deze studie het piloottesten van een methode was, zijn de resultaten dus geen aanbevelingen voor het opnemen of uitsluiten van specifieke aandoeningen in de NBS-programma's. Wel kunnen we uit het toepassen van de methodologie en de eindresultaten een aantal lessen trekken, die werden besproken met de leden van de stuurgroep, en die hieronder worden beschreven, samen met de resultaten.

^b Zes verschillende methodes om de geaggregeerde globale eindscores te berekenen worden gebruikt; één basisanalyse en vijf varianten. Zie wetenschappelijk rapport voor details.



3. SELECTIE EN DEFINITIE VAN DE CRITERIA

Er werden zeven criteria geselecteerd om aandoeningen te scoren, gebaseerd op de INESSS-methodologie.

De definitie van deze criteria werd verder verfijnd in de stuurgroep, om zo goed mogelijk aan te sluiten bij de criteria die reeds in de gemeenschappen, worden gehanteerd, met zo weinig mogelijk overlappingsen. We wilden hiermee vermijden dat dezelfde aspecten van screening dubbel worden gescoord binnen meerdere criteria,. De criteria en de beschrijving in Tabel 3 zijn een verkorte versie van de informatie per criterium. De volledige lijst staat in het wetenschappelijk rapport.



Tabel 3 –Definitie van de criteria voor het scoren van de aandoening

Criteria	Beschrijving
Frequentie (geboorteprevalentie^c)	Geboorteprevalentie in VL, FWB of België, indien beschikbaar, en in West-Europa, Noord-Amerika of wereldwijd
Ernst van de aandoening in onbehandelde gevallen	Ernst van de aandoening zonder vroegtijdige behandeling (natuurlijk verloop), met inbegrip van sequelae, mortaliteit en verlies aan kwaliteit van leven door toedoen van de aandoening
Tijdige beschikbaarheid van de testresultaten	Zijn de testresultaten op tijd beschikbaar om vermijdbare complicaties en sequelae te kunnen voorkomen?
Doeltreffendheid van vroegtijdige behandeling vs. laattijdige behandeling, volgens type behandeling (specifiek en niet-specifiek)	Hoe doeltreffend is vroegtijdige behandeling in vergelijking met laattijdige behandeling? Is er bewijs dat deelnemers (pasgeborenen) waarschijnlijk zullen gebaat zijn bij deze screening en er niet zullen door worden geschaad, bijvoorbeeld door het vervroegen van de diagnose zonder dat dit een noemenswaardige impact heeft op de klinische uitkomst? Graad van consensus over het diagnostische pad na een positief testresultaat Graad van consensus over aanpak van de aandoening indien bevestigd Beschikbaarheid van diagnostische en behandelingsfaciliteiten
Waarschijnlijkheid en impact van vals positieve resultaten	Worden er vals positieve resultaten gemeld: hoe frequent zijn ze en wat is de impact?
Waarschijnlijkheid en impact van vals negatieve resultaten	Worden er vals negatieve resultaten gemeld: hoe frequent zijn ze en wat is de impact?
Impact op het gezondheidszorgsysteem	Impact op diagnostische capaciteit en behandelingscapaciteit, met inbegrip van aanpak van opgespoorde (werkelijke) gevallen en van toevallige vastgestelde diagnoses Kostprijs van (de toevoeging van) een specifieke aandoening aan het bestaande programma Kostprijs van bevestigende diagnostische tests wanneer screening positief blijkt (correct positief en vals positief) Bijkomende kostprijs van een vroeger gestarte behandeling van bevestigde ziekte en kostprijs van een te laat gestarte behandeling (zorgbetaler en maatschappelijk) Kostenefficiëntie van mogelijke screening op deze aandoening (indien beschikbaar) Haalbaarheid van de communicatie naar de deelnemers (ouders) over potentiële voordelen en schade van het screenen van een bijkomende ziekte Aanvaardbaarheid van diagnostische strategie voor de bevolking

^c Geboorteprevalentie: voor definitie zie wetenschappelijk rapport.



Voor twee kwantitatieve criteria werd door de stuurgroep vooraf een bereik van waarden gedefinieerd voor een meer homogene scoring van deze criteria door de beoordeelaars (Tabel 4).

Tabel 4 – Richtlijn scoring voor twee kwantitatieve criteria

Score	Frequentie aandoening: Geboorteprevalentie (per 100 000 levend geboren)	Tijdige beschikbaarheid van testresultaten: Percentage tijdig beschikbare resultaten om ernstige complicaties te voorkomen
1	<1	< 50%
2	1 tot 5	50% tot <80%
3	>5 tot 10	80% tot <100%
4	>10 / 100 000	100%

4. BEPALING VAN HET GEWICHT VAN ELK CRITERIUM

Elk lid van de stuurgroep gaf een gewicht aan elk criterium om het belang voor de beleidsbeslissing weer te geven, met behulp van een LIKERT-schaal op vier punten. De voorlopige resultaten werden daarna besproken in de stuurgroep en nadat enkele criteria waren verduidelijkt en verder gedefinieerd, werden de criteria opnieuw gewogen.

Het gemiddelde en de mediaan van het gewicht dat de beoordeelaars aan elk criterium gaven (zie Figuur 1) vertonen geen uitschieters. Wel gaven ze een iets hoger gewicht aan 'ernst van de aandoening', 'doeltreffendheid van vroegtijdige behandeling vs. laattijdige behandeling' en de 'waarschijnlijkheid en impact van vals negatieve resultaten'.

Deze metingen verbergen echter een heterogeniteit tussen de gewichten gegeven door de beoordeelaars (zie Figuur 2). Het criterium 'ernst van de aandoening' werd unaniem beschouwd als zeer belangrijk, maar er werden erg uiteenlopende antwoorden gegeven voor 'frequentie van de aandoening', 'tijdige beschikbaarheid van de testresultaten' en 'de waarschijnlijkheid en impact van vals negatieve resultaten'. De verschillen in gewichten werden besproken in de stuurgroep. Zij worden waarschijnlijk door de volgende factoren veroorzaakt:

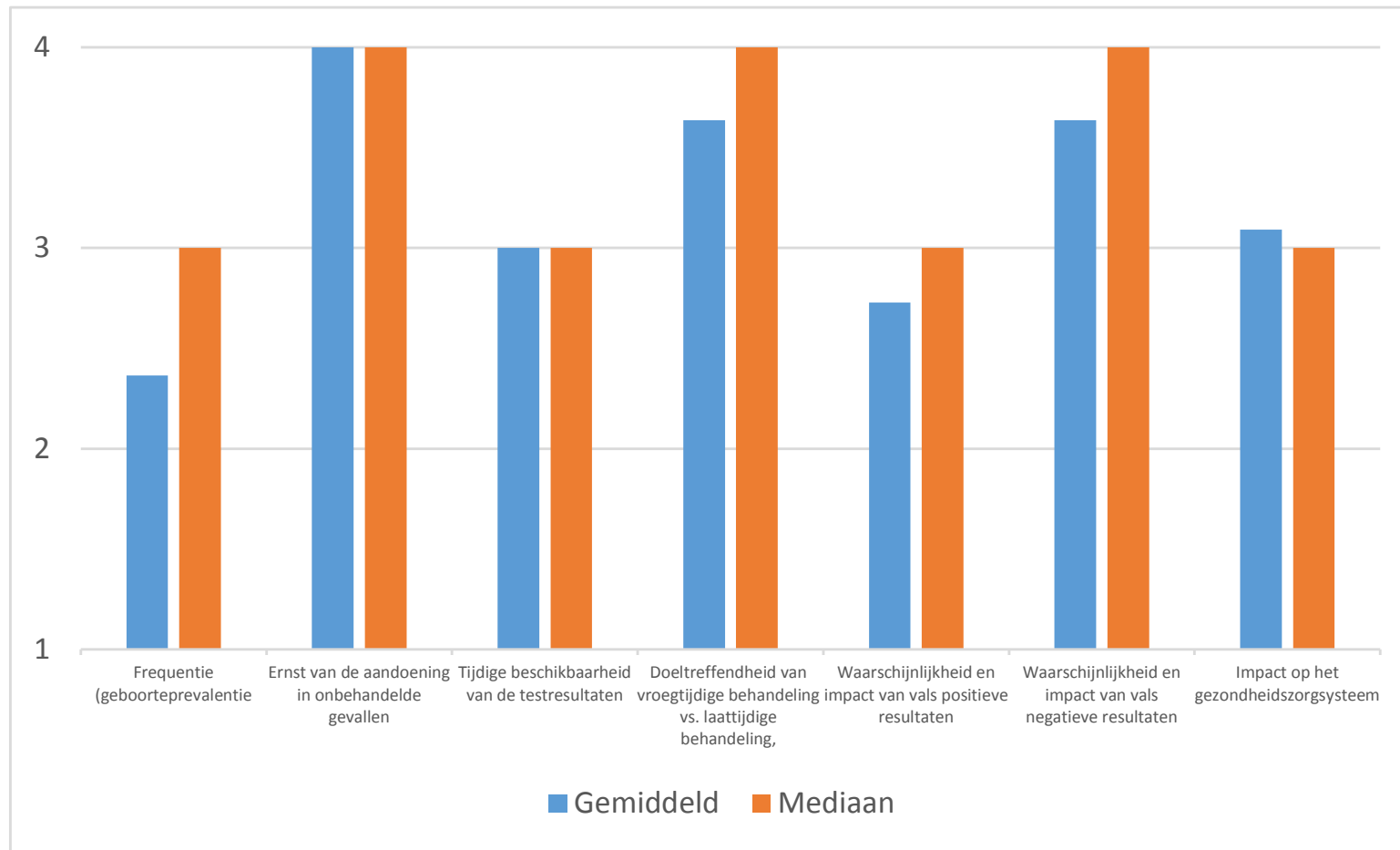
- Hoewel elk criterium werd gedefinieerd en verder in detail uitgewerkt, bleef hun exacte inhoud vaak nog onduidelijk. De leden van de stuurgroep stelden voor om voortaan een nog preciezere beschrijving (bijv. checklist formaat) van de inhoud van elk criterium te geven en om dit ook mondeling toe te lichten aan de beoordeelaars.
- Het profiel van de beoordeelaars beïnvloedde waarschijnlijk het gewicht dat ze aan de criteria toekenden. Beleidsverantwoordelijken, experts uit de NBS-labo's, klinische experts, patiënten vertegenwoordigers of ethische experts kunnen verschillende prioriteiten hebben. Zo was het criterium "tijdige beschikbaarheid van de testresultaten" voor sommige laboratoriumexperts minder belangrijk, omdat zij de snelheid waarmee zij de resultaten meedelen automatisch aanpassen aan de specifieke aandoening (bijvoorbeeld via telefoon i.p.v. met de post). Bij toekomstige prioriteringsoefeningen zal men moeten zorgen voor een evenwichtige verdeling van deze profielen in de beoordelingsgroep.



- Het criterium “impact op het gezondheidszorgsysteem” heeft een erg ruime invulling, met economische en organisatorische haalbaarheids- en aanvaardbaarheidskwesties, waardoor het moeilijk te wegen is (en ook te scoren per ziekte). De inhoud omvat screening en verdere opvolging van opgespoorde gevallen, terwijl dit in België op de twee niveaus van de gezondheidszorg verschillend georganiseerd wordt: screening maakt deel uit van gezondheidspreventie, en dat valt onder de bevoegdheid van de gemeenschappen. Opvolging en behandeling behoren tot de curatieve gezondheidszorg, waarvoor de federale sociale zekerheid (RIZIV) bevoegd is.

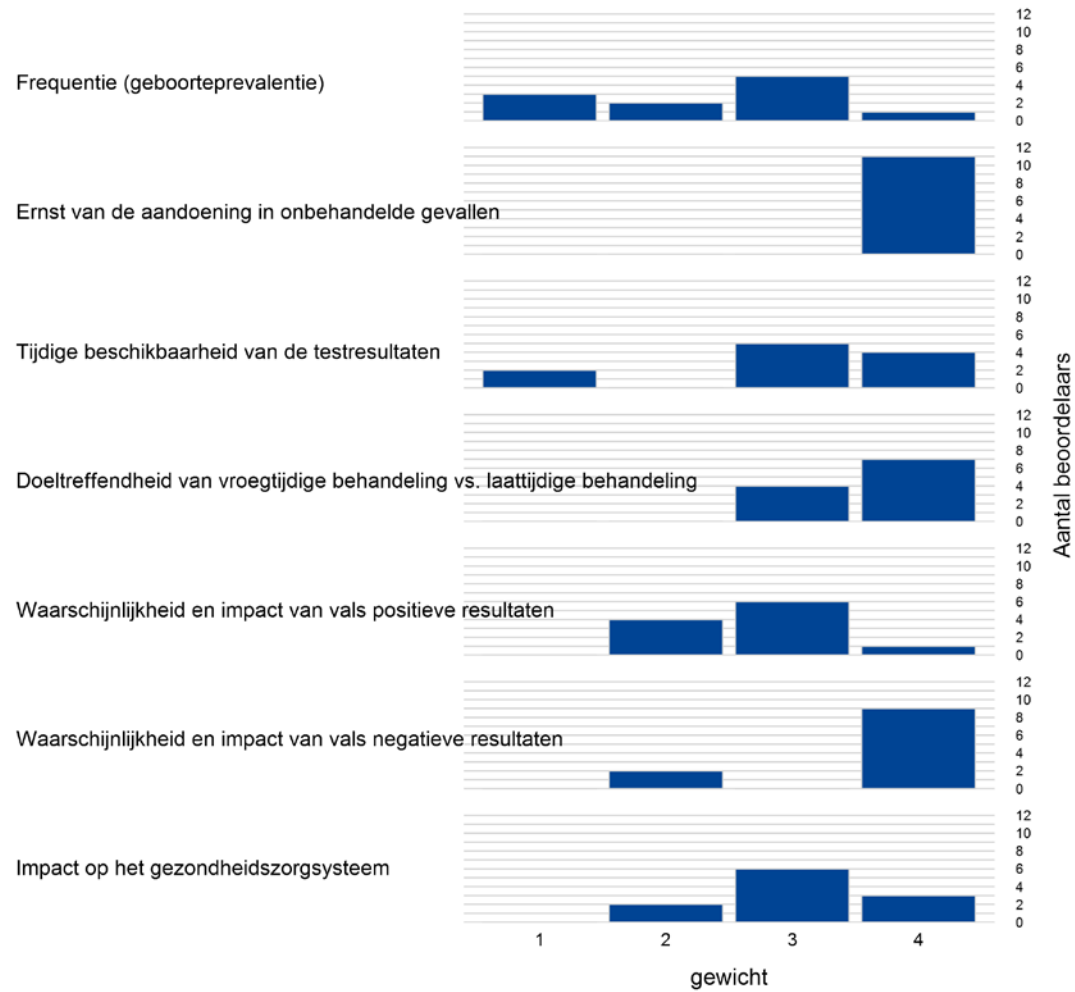


Figuur 1 – Gewicht (gemiddeld en mediaan) voor de zeven criteria (op een LIKERT-schaal van 1 tot 4)





Figuur 2 – Verdeling van de gewichten door de beoordelaars gegeven aan de criteria





5. VERZAMELEN VAN INFORMATIE OVER ELKE AANDOENING

Voor de zes geselecteerde aandoeningen zamelde het KCE via een pragmatisch literatuuronderzoek informatie in over elk van de criteria: Belgische gegevens uit de gemeenschappen en informatie van andere agentschappen (internationaal), de beoordelaars van de stuurgroep en wetenschappelijke en grijze literatuur. Door de lage frequentie van deze aandoeningen was er niet altijd voldoende informatie voorhanden, vooral de informatie over de impact op het gezondheidszorgsysteem was schaars. De informatie werd vervolgens samengevat in een korte ziektefiche en werd bezorgd aan de beoordelaars als gemeenschappelijke en objectieve basis voor de scoring.

6. HET TOEKENNEN VAN DE SCORES PER AANDOENING

De beoordelaars scoorden de zes aandoeningen voor elk van de zeven criteria, wat neerkwam op 42 scores per beoordelaar. De voorlopige scoringsresultaten werden achteraf ter validatie voorgelegd aan de leden van de stuurgroep. Informatie over de criteria en de aandoeningen werd verder verduidelijkt en de beoordelaars konden hun scores aanpassen in geval van misverstanden.

Figuur 3 geeft de gemiddelde en de mediane scores weer per criterium (zonder weging met het gewicht van het criterium) voor de zes aandoeningen. Er zijn een aantal trends: meestal een hoge score (mediaan ≥ 3) voor 'ernst van de aandoening', 'tijdige beschikbaarheid van de testresultaten' en 'doeltreffendheid van vroegtijdige vs. laattijdige behandeling', terwijl 'de waarschijnlijkheid en impact van vals negatieve en positieve resultaten' vaak lage scores haalden, namelijk een mediaan van 1-2 (met uitzondering van Tyrosinemie Type I).

De verdeling van de scores per criterium toont ook grote verschillen tussen de individuele antwoorden, met uitzondering van het criterium 'ernst van de aandoening' (detail gegevens in het wetenschappelijk rapport). De grootste verschillen tussen de scores waren er voor twee criteria: 'waarschijnlijkheid en impact van vals negatieve resultaten' en 'impact op het gezondheidszorgsysteem'. De redenen voor deze discrepanties werden besproken in de stuurgroep. Hieruit kwamen een aantal hypothesen naar voor:

- Een eerste mogelijke verklaring is dat de gegevens onvolledig zijn voor verschillende criteria. Het gaat hier immers om zeldzame aandoeningen. Door een gebrek aan voldoende harde gegevens hadden de persoonlijke kennis en ervaring van de beoordelaars waarschijnlijk een belangrijke invloed op de scores.
- Voor waarschijnlijkheid en impact van vals negatieve resultaten is het mogelijk dat de waarschijnlijkheid van een vals negatief resultaat hoog kan zijn maar dat de impact ervan laag is. Daardoor kunnen deze aspecten moeilijk worden samengevat in één enkele score.
- Het criterium "impact op het gezondheidszorgsysteem" kreeg systematisch uiteenlopende scores (1 tot 4 of 2 tot 4), wat niet wordt

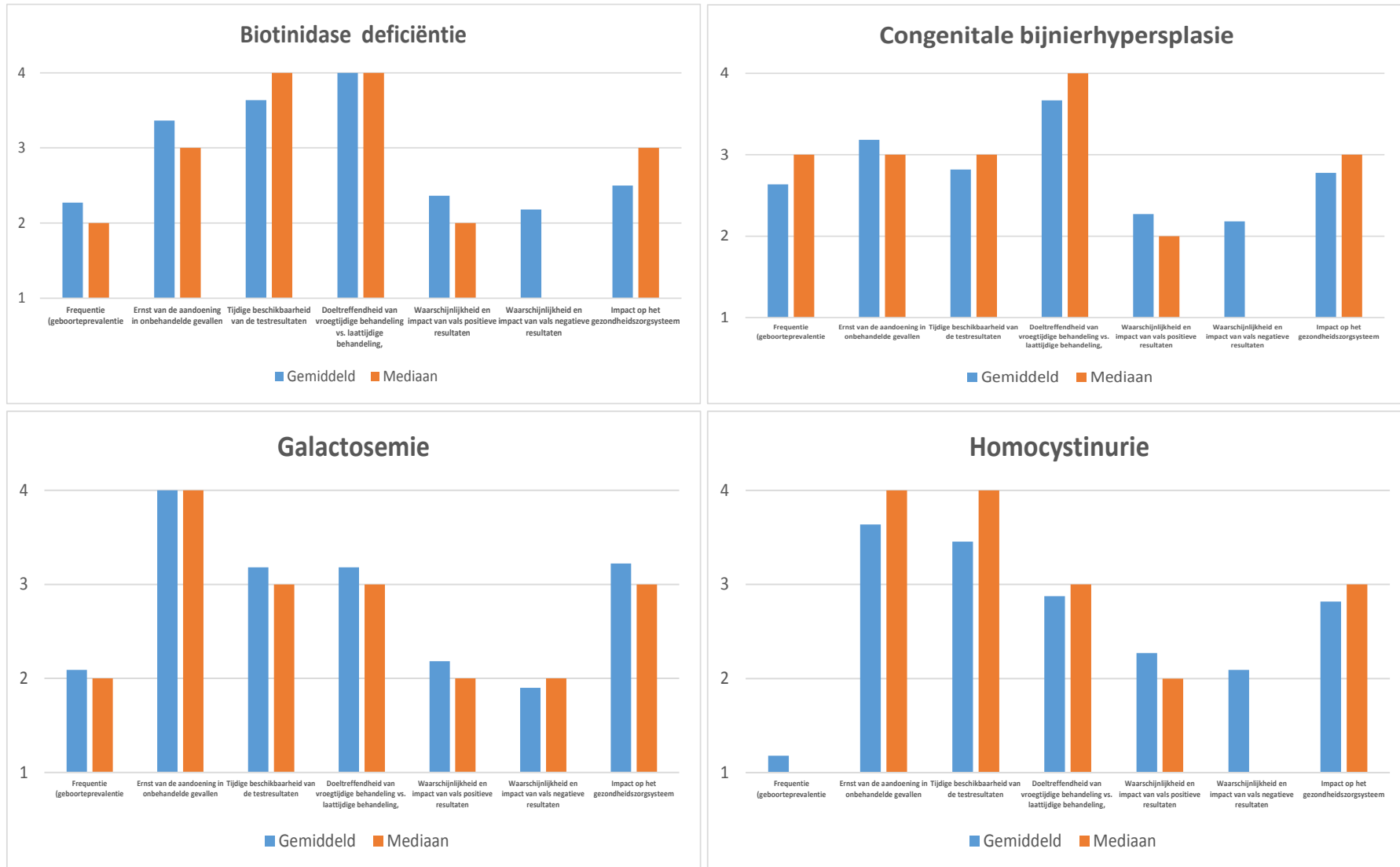


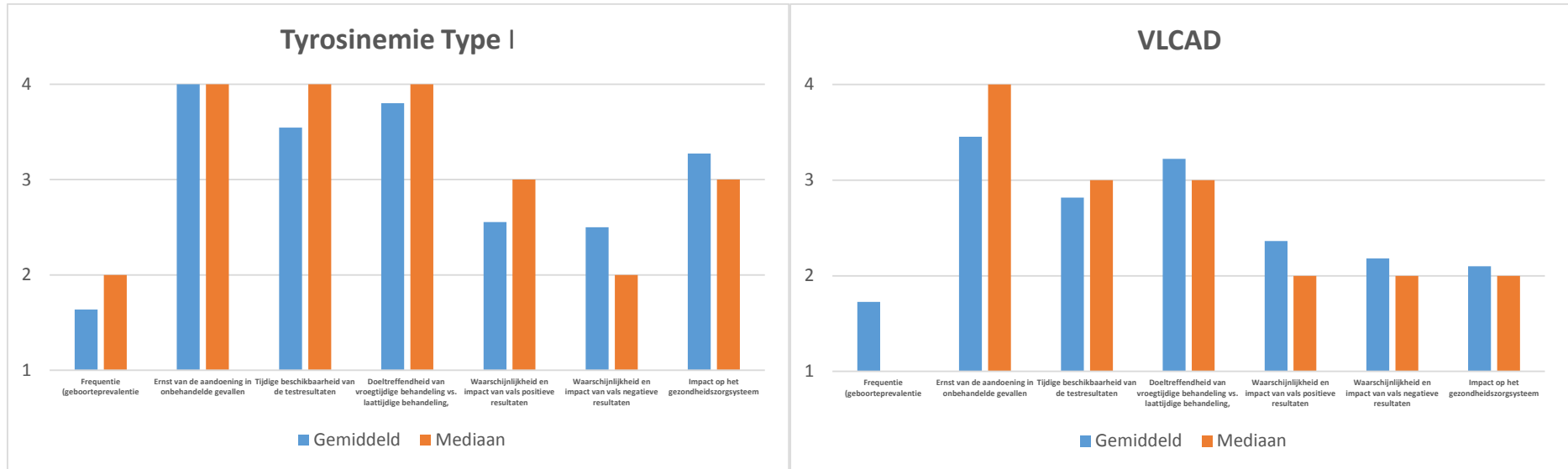
weergegeven in het gemiddelde en de mediaan van de scores (meestal rond 3). Dit komt door de brede definitie van dit criterium, dat uiteenlopende componenten bevat (Tabel 3), en door de verschillende profielen van de beoordelaars. Zo waren beoordelaars die betrokken zijn bij beleidsvorming gevoeliger voor economische en budgetkwesties dan laboratoriumexperten, die meer belang hechtten aan de kenmerken van de screeningtest.

- Beleidsverantwoordelijken van de gemeenschappen en van de NBS-laboratoria kunnen meer belang hechten aan de impact op vroegtijdige detectie, terwijl anderen gevoeliger zijn voor de impact op de curatieve gezondheidszorg.



Figuur 3 – Gemiddelde en mediaan van de scores voor elk criterium per aandoening (niet gewogen voor het gewicht van het criterium)





VLCAD: *Very long-chain acyl-CoA dehydrogenase*



7. GLOBALE EINDSCORE PER ZIEKTE

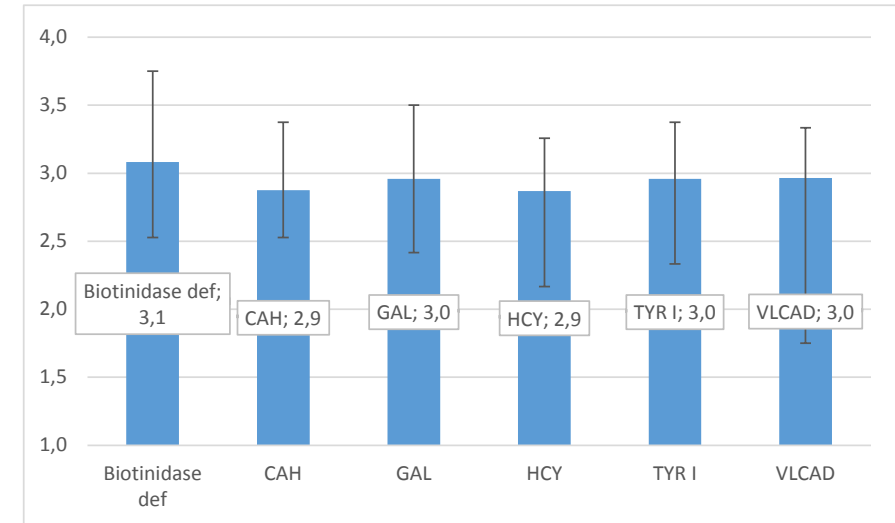
De meest opvallende vaststelling van deze MCDA-piloottest is dat in de basisanalyse de globale eindscores voor de zes ziekten weinig verschil vertonen, waardoor het moeilijk is om ze te rangschikken naar prioriteit (Figuur 4). De mediane waarden van deze globale eindscores gaan van 2.9 tot 3.1 maar met een grote spreiding voor de individuele scores over beoordelaars heen.

Zoals besproken in de methodologie bekijken we ook nog vijf andere methodes om de globale eindscores per aandoening te berekenen en te vergelijken (zie scope en methodologie). Figuur 5 toont aan dat de rangschikking van elke aandoening enigszins afhankelijk is van de methode om de globale eindscores te berekenen. Er zijn echter bepaalde patronen; biotinidase deficiëntie krijgt bij alle methodes de eerste of tweede plaats, GAL krijgt de eerste of tweede plaats bij drie van de zes methodes en HCY en VLCAD krijgen bijna systematisch een lage rangschikking.^d

De mogelijke verklaringen voor deze kleine verschillen tussen de globale eindscores werden besproken in de stuurgroep:

- De zes aandoeningen hebben waarschijnlijk een gelijkaardig belang voor NBS-screening in België, omdat ze al in één gemeenschap worden gescreend, en dus belangrijk genoeg bevonden werden.
- De smalle scoreschaal van 1 tot 4 kan potentiële verschillen tussen de individuele scores kleiner maken. Een schaal van 1 tot 10 kan andere resultaten opleveren.
- De gebruikte indicatoren zijn de mediaan of het gemiddelde (“centrale tendens” metingen), die de verschillen tussen individuele scores verbergen.

Figuur 4 – Basisanalyse: globale eindscores per aandoening (en spreiding van individuele scores over de beoordelaars heen)

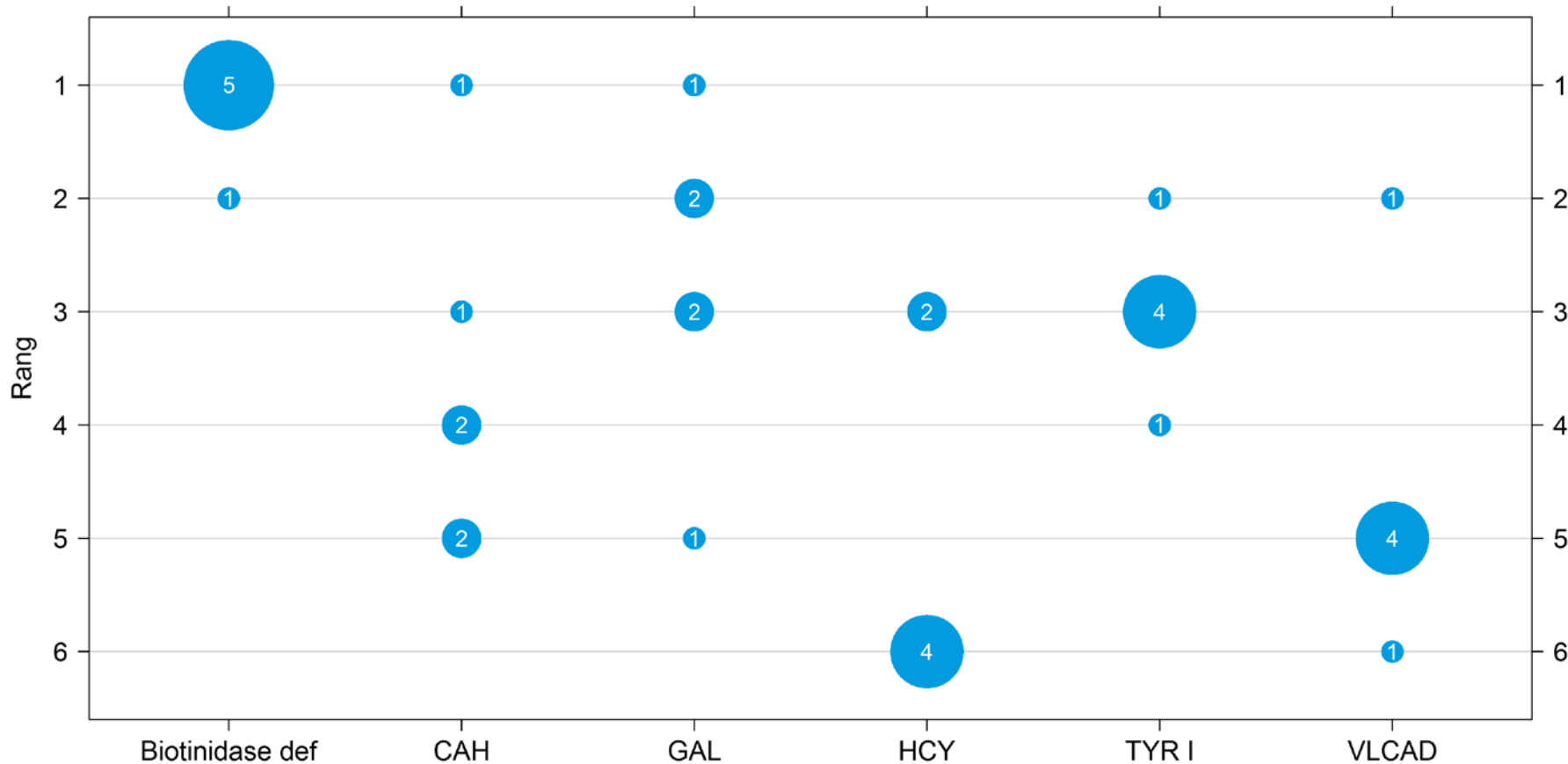


Biotinidase def: Biotinidase deficiëntie; CAH: Congenitale bijnierhyperplasie; GAL: Galactosemie; HCY: Homocystinurie; TYR I: Tyrosinemie Type I; VLCAD: Very long-chain acyl-CoA dehydrogenase

^d Als gevolg van enkele ex aequo 's is de som van elk van de rangschikkingen niet precies zes.



Figuur 5 – Bijkomende analyses: verdeling van de rangschikking per aandoening voor de zes verschillende methodes om globale eindscores te berekenen



Biotinidase def: Biotinidase deficiëntie; CAH: Congenitale bijnierhyperplasie; GAL: Galactosemie; HCY: Homocystinurie; TYR I: Tyrosinemie Type I; VLCAD: Very long-chain acyl-CoA dehydrogenase



8. CONCLUSIE

In deze studie testen we een MCDA-methode voor de prioritering van aandoeningen die door neonatale bloedscreening kunnen opgespoord worden. Het vinden van een consensus over een selectie en een uniforme definitie van criteria bleek complexer dan verwacht, en vereiste behoorlijk wat tijd en overleg. In deze studie vinden we weinig verschil tussen de globale eindscores om de ziekten te kunnen rangschikken voor NBS.

De rangschikking verschilt ook enigszins, afhankelijk van de methode voor het berekenen van de globale eindscores. Dit maakt een verdere discussie over de methode noodzakelijk.

Verschillende leden van de stuurgroep vonden het ook niet gepast om het beleid en de beslissingen over een dergelijke complexe materie enkel te baseren op een globale eindscore per ziekte. Maar omdat het wegen en scoren voor meer reflectie zorgt en aanzet tot het expliciet maken van argumenten, vond de stuurgroep de resultaten van deze oefening nuttig ter ondersteuning van de beleidsvorming.

Deze methode kan worden verbeterd door elk criterium beter te definiëren, door criteria met tegengestelde componenten op te splitsen, door eventueel een ruimere schaal om te wegen en te scoren te overwegen, en door zorgvuldig na te denken over de profielen van de beoordelaars die nodig zijn om tot een evenwichtige evaluatie te komen.

Toekomstige oefeningen bevatten best meer aandoeningen en kunnen gebruik maken van de lessen die we uit de huidige piloottest kunnen trekken.

9. KERNBOODSCHAPPEN EN LESSEN VOOR DE TOEKOMST

- De besluitvorming over het al dan niet opnemen van een (bijkomende) aandoening in een neonataal screeningsprogramma is een complex proces met verschillende dimensies. Voor dergelijke complexe beslissingsprocessen wordt steeds meer gebruik gemaakt van formele beoordelingsmethoden.
- Steunend op de ervaring van het INESSS in Québec, werd een pilootstudie uitgevoerd om de bruikbaarheid van *Multi-Criteria Decision Analysis (MCDA)* na te gaan. Samen met een stuurgroep van vertegenwoordigers en experts van de gemeenschappen en van stakeholders, besloten we het instrument uit te testen op de zes aandoeningen die momenteel enkel in één van beide gemeenschappen worden opgespoord.
- Het selecteren van de beoordelingscriteria is niet eenvoudig. De soms erg heterogene scores voor het wegen van bepaalde criteria wijst erop dat men in dit proces best meerdere cycli doorloopt, waarbij de criteria verder gespecificeerd en, indien nodig, aangevuld of opgesplitst worden.
- Beoordelaars nemen soms erg uiteenlopende standpunten in. Het toekennen van een consensus gewicht en score aan elk van de weerhouden criteria vereist een uitgebreide discussie tussen de beoordelaars.
- Het toekennen van een score voor elk criterium voor alle ziekten, vereist het voorafgaandelijk opzoeken van de beschikbare evidentie rond de epidemiologie, rond beschikbare tests en interventies, met inbegrip van de organisatorische, gezondheid economische en ethische aspecten. Deze informatie moet aan alle beoordelaars worden overgemaakt, zodat ze allen met dezelfde basisinformatie kunnen scoren.
- Voor het bepalen van de gewichten per criterium en voor het scoren van elke ziekte voor dit criterium, is een LIKERT schaal met slechts vier punten mogelijk te weinig discriminerend. We konden hier echter geen uitsluitel geven. De zes ziekten bevonden zich wellicht in dezelfde zone tussen verantwoord binnen een screeningsprogramma of niet, omdat zij slechts door één gemeenschap werden geselecteerd.



- Bij het samenstellen van het panel beoordelaars moeten alle relevante standpunten vertegenwoordigd zijn, zodat geen essentiële argumenten over het hoofd worden gezien.
- Tijdens de gesprekken binnen de stuurgroep werd opgemerkt dat het misschien nuttig zou kunnen zijn te overwegen om in de schaal voor het scoren van de ziekten een 'uitsluitingsscore' te voorzien, die aangeeft dat een noodzakelijke voorwaarde voor screening voor deze ziekte hoe dan ook niet vervuld is.
- De manier waarop de globale eindscores per ziekte worden berekend, kan de uiteindelijke rangschikking van de ziekten beïnvloeden. Ook hier konden met de zes geselecteerde aandoeningen geen verdere conclusies getrokken worden.
- Het is niet de bedoeling dat de beslissing rond het opnemen of uitsluiten van ziekten in een screeningsprogramma louter op basis van een MCDA rangschikking wordt genomen. Een dergelijke oefening laat wel toe om de discussies rond een besluit op een objectievere en transparantere manier te laten verlopen. Zo kan er een grotere coherentie worden bereikt tussen opeenvolgende beslissingen.



■ AANBEVELINGEN^e

Aan de beleidsverantwoordelijken van de deelstaten

- Er is internationaal geen consensus over welke aandoeningen dienen opgespoord te worden binnen een neonataal bloedscreeningsprogramma, noch over hoe deze beslissingen moeten genomen worden. Voor deze complexe vragen kan een Multi-Criteria Decision Analysis methode (MCDA) een bijkomend hulpmiddel zijn bij de besluitvorming.
- MCDA kan niet als een louter mathematische beslissingsvorm gebruikt worden maar het kan helpen de beslissingsprocessen te structureren en transparanter te maken. Het systematisch toepassen van een dergelijke methode kan op termijn ook leiden tot meer consistentie in de besluitvorming.
- De criteria die gebruikt worden moeten vooraf grondig besproken en gedefiniëerd worden met de beleidsverantwoordelijken en de volledige groep van de beoordelaars. Mogelijks moeten sommige criteria die zowel gerelateerd zijn aan de ernst van de aandoening als aan de mogelijke behandeling verder uit elkaar worden gehaald omdat deze tegengesteld kunnen zijn.
- Er moet verder onderzocht worden wat de beste samenvattende uitkomstmaat is voor deze vorm van MCDA.
- Voor elke ziekte die beoordeeld wordt, moet, per criterium, de nodige informatie verzameld worden en ter beschikking van de beoordelaars worden gesteld.
- Bij de samenstelling van de groep beoordelaars moet aandacht gaan naar een voldoende evenwicht tussen de verschillende betrokken partijen (beleidsverantwoordelijken, specialisten, patiëntenvertegenwoordigers, ethici, etc...) om alle opinies en voorkeuren zo goed mogelijk te vertegenwoordigen.
- Informatie over de kosten en baten van het diagnostisch traject na een positieve screeningstest (inclusief de vals positieven) en van de uiteindelijke behandeling van de terecht positieven blijkt moeilijk te verkrijgen. We bevelen aan om een systematische en langdurige opvolging te voorzien van alle positief gescreende neonati om hierop in de toekomst een beter antwoord te vinden.

^e Het KCE is als enige verantwoordelijk voor de aanbevelingen.



COLOFON

- Titel:** Multicriteria beslissingsanalyse bij de selectie van prioritair aandoeningen binnen de neonatale bloedscreening – Synthese
- Auteurs:** Chris De Laet (KCE), Germaine Hanquet (KCE), Erik Hendrickx (KCE)
- Project coordinator:** Marijke Eyssen (KCE).
- Reviewers:** Gudrun Briat (KCE), Irina Cleemput (KCE), Frank Hulstaert (KCE), Raf Mertens (KCE), Karin Rondia (KCE), Genevieve Veereman (KCE).
- Externe experts:** François Boemer (Université de Liège), Marc Boëne (BOKS, Belgische Organisatie voor Kinderen en volwassenen met een Stofwisselingsziekte / Organisation Belge pour des enfants et des adultes souffrant d'une maladie métabolique), Vincent Bours (Université de Liège), Karen Colaert (Vlaams Agentschap Zorg en Gezondheid), Lutgarde De Baere (BOKS, Belgische Organisatie voor Kinderen en volwassenen met een Stofwisselingsziekte / Organisation Belge pour des enfants et des adultes souffrant d'une maladie métabolique), Linda De Meirleir (AZ-VUB), François Eyskens (Provinciaal Centrum voor de opsporing van metabole aandoeningen), Philippe Goyens (CHU Brugmann en HUDERF), Hilde Laeremans (AZ-VUB en CHU Brugmann), Marie-Christine Mauroy (Office de la Naissance et de l'Enfance (ONE)), Heidi Mertes (Ugent), Marie-Cécile Nassogne (Cliniques universitaires St Luc), Tatiana Pereira (Fédération Wallonie Bruxelles (FWB) en ONE), Roland Schoos (Université de Liège), Elfriede Swinnen (Orphanet), Béatrice Toussaint (SPF Santé Publique / FOD Volksgezondheid), Pieter Vandenbulcke (Vlaams Agentschap Zorg en Gezondheid), Marie-Françoise Vincent (Cliniques universitaires St Luc), Véronique Zinnen (FWB en ONE).
- Externe validatoren:** Jean-Jacques Cassiman (KU Leuven), Corinne De Laet (HUDERF, Brussels), Brigitte Côté (Institut National d'Excellence en Santé et en Services Sociaux (Québec, Canada)).
- Acknowledgements:** Muriel Deguerry (Observatorium voor Gezondheid en Welzijn van Brussel-Hoofdstad) voor de specifieke informatie over de screening in de regio Brussel – hoofdstad. Stephan Devriese (KCE) voor zijn hulp bij het maken van specifieke grafieken. Gudrun Briat (KCE) en Karin Rondia (KCE) voor nalezen en verbeteren van de synthese in het Nederlands en het Frans. Cover foto's: De Kern – verloskundigenpraktijk (NL) en Andrée Mangin (KCE).
- Andere gemelde belangen:** Geen gemeld François Boemer (Université de Liège), Jean-Jacques Cassiman (KU Leuven), Corinne De Laet (HUDERF), Linda De Meirleir (AZ-VUB), François Eyskens (Provinciaal Centrum voor de opsporing van metabole aandoeningen), Heidi Mertes (Ugent), Marie-Françoise Vincent (Cliniques universitaires St Luc) et Véronique Zinnen (FWB et ONE).
- Lidmaatschap van een belangengroep op wie de resultaten van dit rapport een impact kunnen hebben: Hilde Laeremans (ULB en Vlaams Centrum Brussel voor opsporing van aangeboren Metabole Aandoeningen), Tatiana Pereira (FWB, Direction Générale de la Santé Exécutant pour ONE) en Roland Schoos (Université de Liège, Centre de dépistage néonatal de Liège).



Voorzitterschap of verantwoordelijke functie in een instelling, vereniging, afdeling of andere entiteit waarop de resultaten van dit rapport een impact kunnen hebben: Karen Colaert (organiseert het screeningsprogramma “aangeboren aandoeningen bij pasgeborenen via een bloedstaal” bij het Vlaams Agentschap Zorg en Gezondheid) Lutgarde De Baere (voorzitter van BOKS een patiënten organisatie voor metabole ziekten), Philippe Goyens en Hilde Laeremans (coördinatoren van een laboratorium voor de screening van of metabole ziekten voor ULB en AZ-VUB (VCBMA), Roland Schoos (lid van de stuurgroep voor neonatale screening van metabole ziekten van de FWB), Béatrice Toussaint (verantwoordelijk voor neonatale screening bij het cabinet van de FWB in 2009-2014), Pieter Vandenbulcke (hoofd van het team algemene preventie bij het Vlaams Agentschap Zorg en Gezondheid).en voor bevolkingsonderzoek programma’s gerelateerd aan ziekte preventie.

Andere mogelijke belangen die kunnen leiden tot een potentieel of reëel belangenconflict: Brigitte Côté (was principal investigator voor het wetenschappelijke INESSS onderzoek (2013) “Pertinence d’élargir le programme de dépistage néonatal sanguin au Québec”).

Layout:

Joyce Grijseels, Sophie Vaes.

Disclaimer:

- **De externe experts werden geraadpleegd over een (preliminaire) versie van het wetenschappelijke rapport. Hun opmerkingen werden tijdens vergaderingen besproken. Zij zijn geen coauteur van het wetenschappelijke rapport en gingen niet noodzakelijk akkoord met de inhoud ervan.**
- **Vervolgens werd een (finale) versie aan de validatoren voorgelegd. De validatie van het rapport volgt uit een consensus of een meerderheidsstem tussen de validatoren. Zij zijn geen coauteur van het wetenschappelijke rapport en gingen niet noodzakelijk alle drie akkoord met de inhoud ervan.**
- **Tot slot werd dit rapport unaniem goedgekeurd door de Raad van Bestuur (zie <http://kce.fgov.be/nl/content/de-raad-van-bestuur>).**
- **Alleen het KCE is verantwoordelijk voor de eventuele resterende vergissingen of onvolledigheden alsook voor de aanbevelingen aan de overheid.**

Publicatiedatum: 25 april 2016
Domein: Health Services Research (HSR)
MeSH: Decision Making; Metabolism, Inborn Errors; Neonatal Screening
NLM classificatie: WD 205
Taal: Nederlands
Formaat: Adobe® PDF™ (A4)
Wettelijk depot: D/2016/10.273/41
ISSN: 2466-6432



Copyright:

De KCE-rapporten worden gepubliceerd onder de Licentie Creative Commons « by/nc/nd » <http://kce.fgov.be/nl/content/de-copyrights-van-de-kce-rapporten>.



Hoe refereren naar dit document?

De Laet C, Hanquet G, Hendrickx E. Multicriteria beslissingsanalyse bij de selectie van prioritaire aandoeningen binnen de neonatale bloedscreening – Synthese. Health Services Research (HSR). Brussel: Federaal Kenniscentrum voor de Gezondheidszorg (KCE). 2016. KCE Reports 267As. D/2016/10.273/41.

Dit document is beschikbaar op de website van het Federaal Kenniscentrum voor de Gezondheidszorg.