

## SYNTHÈSE

# DÉPISTAGE NÉONATAL SANGUIN : ANALYSE DE DÉCISION MULTICRITÈRE POUR SÉLECTIONNER LES MALADIES PRIORITAIRES





## SYNTHÈSE

# DÉPISTAGE NÉONATAL SANGUIN : ANALYSE DE DÉCISION MULTICRITÈRE POUR SÉLECTIONNER LES MALADIES PRIORITAIRES

CHRIS DE LAET, GERMAINE HANQUET, ERIK HENDRICKX





## ■ PRÉFACE

En l'an 2016, en Europe, nous pourrions raisonnablement espérer que les soins de santé aient peu à peu atteint un certain degré d'harmonisation. N'avons-nous pas, pour chaque discipline, des 'Sociétés européennes' qui produisent consciencieusement des recommandations de bonnes pratiques et des consensus, théoriquement tous basés sur les mêmes sources internationales *evidence-based* ? Pourtant, en ce qui concerne le dépistage des maladies métaboliques chez les nouveau-nés, rien n'est moins vrai. Selon le pays où le petit Européen verra le jour, sa piqûre au talon portera sur une, deux... ou trente analyses, pour dépister des affections chaque fois différentes. Il est vrai que tous les Européens ne sont pas issus de pools génétiques comparables : il n'y aurait sans doute pas beaucoup de sens à dépister la thalassémie en Laponie... encore que, avec les vagues de migrations que nous connaissons, la situation pourrait bien se compliquer. Mais même entre pays – ou régions – qui dépistent un nombre comparable de maladies, la composition des panels d'analyses peut varier considérablement.

Le choix d'un panel d'analyses 'optimal' est bien entendu une question complexe, dans laquelle doivent être mises en balance toute une série de dimensions. Et nous connaissons tous les réunions de commission où il faut démêler ce genre d'écheveaux complexes : les discussions qui sautent sans cesse d'une dimension du problème à l'autre, les leaders d'opinion qui monopolisent le débat, chacune des parties prenantes qui tente de faire valoir son propre éclairage, à tel point qu'en fin de compte, personne ne peut garder une vue globale de la question. Et cela se termine invariablement par des listes différentes pour chaque pays ou région, toujours avec de bonnes raisons, mais souvent contingentes et peu consistantes.

Ici et là, l'idée fait son chemin que des décisions comme celles sur le dépistage néonatal pourraient se dérouler de manière un peu plus cohérente, et sur ce point, nous avons trouvé l'exemple des Québécois très inspirant. Dans l'étude que vous avez devant vous, nous avons donc tenté d'évaluer dans quelle mesure leur modèle d'analyse de décision multi-critère (MCDA) était applicable à la Belgique (lisez : aux réalités wallonne, flamande et bruxelloise). C'est la deuxième fois que nous pratiquons cet exercice dans le cadre du système de santé belge ; en effet, nous sommes également occupés à tester un modèle MCDA avec l'INAMI dans le dossier des besoins médicaux non rencontrés (*unmet medical needs*).

Nous pouvons espérer y voir les prémices d'un tournant culturel radical qui osera, avec l'aide de tels outils, rendre les prises de décisions dans des domaines aussi complexes beaucoup plus objectives et surtout beaucoup plus transparentes.

Christian LÉONARD  
Directeur Général Adjoint

Raf MERTENS  
Directeur Général



## ■ SYNTHÈSE

### TABLE DES MATIÈRES

■	PRÉFACE.....	1
■	SYNTHÈSE .....	2
1.	PROGRAMMES DE DÉPISTAGE NÉONATAL SANGUIN (DNS) .....	3
2.	CHAMP ET MÉTHODOLOGIE .....	5
3.	SÉLECTION ET DÉFINITION DES CRITÈRES .....	6
4.	DÉTERMINATION DU POIDS DE CHAQUE CRITÈRE .....	8
5.	COLLECTE D'INFORMATIONS SUR CHAQUE PATHOLOGIE .....	12
6.	ATTRIBUTION DES SCORES PAR PATHOLOGIE .....	12
7.	SCORE GLOBAL PAR MALADIE .....	16
8.	CONCLUSION .....	18
9.	MESSAGES-CLÉS ET ENSEIGNEMENTS POUR L'AVENIR .....	18
■	RECOMMANDATIONS.....	20



## 1. PROGRAMMES DE DÉPISTAGE NÉONATAL SANGUIN (DNS)

Le programme de dépistage néonatal sanguin (DNS) est un programme de santé publique qui soumet systématiquement tous les nouveau-nés, peu après la naissance, au dépistage d'une série de maladies congénitales. La plupart de ces affections sont des maladies métaboliques dont les premiers symptômes ne sont pas détectables à la naissance, mais qui peuvent donner rapidement des séquelles irréversibles si une prise en charge adéquate n'est pas mise en route immédiatement. Les dépister avant l'apparition de ces premiers signes cliniques permet de poser un diagnostic précoce et d'intervenir en temps utile. Également appelé « test du buvard » ou test de Guthrie, le test de dépistage néonatal consiste à prélever chez le nouveau-né un tout petit échantillon de sang qui sera recueilli sur une carte pré-imprimée (dite carte de Guthrie).

Il n'existe pas de consensus international quant aux maladies congénitales à dépister, ce qui explique que leur nombre (de zéro à plus d'une quarantaine) et leur nature varient sensiblement d'un pays à l'autre.

En Belgique, le programme de DNS a débuté en 1968 avec le dépistage systématique de la phénylcétonurie chez tous les nouveau-nés. Par la suite s'y sont progressivement ajoutées six autres maladies.

Au début des années 1980, la responsabilité du programme a été transférée à la Communauté française (rebaptisée plus tard Fédération Wallonie-Bruxelles (FWB))<sup>a</sup> et à la Vlaamse Gemeenschap (VG). Depuis lors, chaque Communauté décide elle-même des maladies qu'elle dépiste en se basant sur ses critères juridiques et sur les recommandations de ses comités de pilotage. De ce fait, la liste des maladies dépistées est différente dans le

Nord et le Sud du pays : elle comporte 13 pathologies en FWB et 11 en VG, dont 9 sont communes aux deux Communautés (Tableau 1).

Les deux Communautés utilisent des technologies comparables, mais les analyses sont effectuées dans des centres de référence distincts (trois en FWB et deux en VG). Pour la majorité des maladies métaboliques congénitales (mais pas toutes) ces analyses sont actuellement réalisées par spectrométrie de masse en tandem (SM/SM) – une technique de laboratoire qui, pour des volumes suffisants, est relativement bon marché et permet de dépister rapidement et simultanément un grand nombre de maladies métaboliques au départ d'un échantillon sanguin unique. Cette technique consiste à mesurer la concentration d'un grand nombre de métabolites ; si le résultat est positif, d'autres tests doivent être réalisés pour confirmation avant qu'un diagnostic final puisse être posé. D'après les chiffres de la VG, le coût total du dépistage primaire (y compris les cartes, l'envoi, la communication des résultats et le matériel de promotion) s'élève à environ €20 par nourrisson

La technique SM/SM permet potentiellement d'identifier un grand nombre d'affections. Le choix de celles qui seront incluses dans les programmes de dépistage se base donc sur la recherche d'un équilibre acceptable entre les bénéfices et les risques. Le bénéfice potentiel est que la détection précoce, suivie d'une intervention efficace, permet d'empêcher le développement de la maladie, l'apparition de ses séquelles et parfois même le décès prématuré. Quant aux risques, ils sont principalement liés aux résultats faux négatifs ou faux positifs, qui peuvent générer respectivement un soulagement trompeur ou une inquiétude, voire des coûts inutiles.

L'objectif de cette étude du KCE est d'aider à la prise de décision sur les pathologies qui pourraient être incluses dans les programmes de DNS menés par les Communautés.

<sup>a</sup> En FWB, le programme DNS est géré par l'Office de la Naissance et de l'Enfance (ONE) depuis 2015.



Tableau 1 – Maladies incluses dans les programmes de dépistage néonatal sanguin des Communautés (février 2016)

Maladie	Abréviation	Vlaamse Gemeenschap	Fédération Wallonie- Bruxelles	Reprise dans cette étude-pilote
<i>Maladies métaboliques</i>				
Déficit en biotinidase	LMCD	Oui	Non	Oui
Galactosémie	GAL	Non	Oui	Oui
Acidémie glutarique de type 1	GA I	Oui	Oui	
Homocystinurie	HCY	Non	Oui	Oui
Acidémie isovalérique	IVA	Oui	Oui	
Maladie du sirop d'érable	MSUD	Oui	Oui	
Déficit en acyl-CoA déshydrogénase des acides gras à chaîne moyenne	MCAD	Oui	Oui	
Acidémie méthylmalonique	MMA	Oui	Oui	
Déficit en multiple acyl-CoA déshydrogénase	MADD	Oui	Oui	
Phénylcétonurie	PKU	Oui	Oui	
Acidémie propionique	PA	Oui	Oui	
Tyrosinémie de type I	TYR I	Non	Oui	Oui
Déficit en acyl-CoA déshydrogénase des acides gras à très longue chaîne	VLCAD	Non	Oui	Oui
<i>Troubles endocriniens</i>				
Hypothyroïdie congénitale	CHT	Oui	Oui	
Hyperplasie congénitale des surrénales	CAH	Oui	Non	Oui





## 2. CHAMP ET MÉTHODOLOGIE

Dans le cadre de cette étude, nous avons mené une étude pilote utilisant une méthode d'analyse décisionnelle multicritère (*Multi-Criteria Decision Analysis* ou MCDA) visant à définir les priorités en termes de maladies à inclure dans un DNS. Cette approche a été appliquée en 2013 au Québec par l'Institut National d'Excellence en Santé et en Services Sociaux (INESSS).

Dans la méthode MCDA, on sélectionne a priori un certain nombre de critères, auxquels on attribue une pondération spécifique. Les maladies évaluées reçoivent ensuite un score pour chacun de ces critères, avant d'être classées sur la base du score global ainsi obtenu.

Notre étude pilote comprend les étapes suivantes :

1. Sélection et description précise des critères d'évaluation des maladies.
2. Pondération de chaque critère pour le calcul du score global de chaque maladie. Cette pondération est exprimée au moyen d'un score de 1 à 4 sur une échelle de LIKERT (Tableau 2).
3. Préparation de courtes fiches synthétiques pour chaque maladie, résumant les informations scientifiques relatives à chaque critère pour cette maladie.
4. Détermination d'un score pour chaque maladie et chaque critère, également exprimé sur une échelle de LIKERT allant de 1 à 4 (Tableau 2).
5. Calcul d'un score global par maladie: en général, ce calcul se fait en multipliant la pondération d'un critère par le score attribué à la maladie pour ce critère, et ces scores pondérés sont alors additionnés pour donner un score par maladie. Ceci donne ou bien directement un score global ou un score par maladie pour chaque évaluateur. Dans ce dernier cas, les scores des différents évaluateurs sont ensuite agrégés pour obtenir un score global par maladie, ce qui permet leur classement.<sup>b</sup>

<sup>b</sup> Six méthodes différentes pour calculer les scores globaux sont utilisées ; une analyse de base et cinq variantes. Voir le rapport scientifique pour les détails.

**Tableau 2 – Description de l'échelle de LIKERT utilisée pour pondérer les différents critères et attribuer un score à la maladie**

Échelle	Poids du critère	Score de la maladie pour ce critère
1	non pertinent	très faible
2	peu pertinent	plutôt faible
3	pertinent	plutôt élevé
4	très pertinent	très élevé
0		impossible à définir, données manquantes

La 1<sup>re</sup> étape a été réalisée par l'équipe du KCE en collaboration avec un groupe de pilotage composé de responsables des deux Communautés, d'experts issus des laboratoires de DNS, d'experts cliniques, de représentants des patients et d'experts en éthique (11 évaluateurs au total). Les étapes 2 et 4 n'ont été réalisées que par les membres du groupe de pilotage, tandis que les étapes 3 et 5 ont été réalisées par l'équipe du KCE puis discutées au sein du groupe de pilotage.

Pour cette étude pilote, le groupe de pilotage et l'équipe du KCE ont décidé d'appliquer la méthodologie aux six pathologies qui ne sont dépistées que dans l'une ou l'autre des Communautés (Tableau 1). Comme l'objectif de cette étude est de tester cette méthodologie, ses résultats ne sont donc pas des recommandations d'inclusion ou d'exclusion d'une maladie spécifique dans les programmes de DNS. Nous pouvons néanmoins tirer de l'application de la méthode choisie et des résultats obtenus un certain nombre d'enseignements, qui ont fait l'objet d'une discussion avec les membres du groupe de pilotage et qui sont décrites ci-dessous avec les résultats.



### 3. SÉLECTION ET DÉFINITION DES CRITÈRES

Sept critères ont été sélectionnés pour attribuer un score aux maladies, sur la base de la méthodologie INESSS.

La définition de ces critères a été affinée au sein du groupe de pilotage pour correspondre le mieux possible aux critères définis par les Communautés, tout en limitant au maximum la duplication de leur contenu. Nous avons voulu éviter par-là que les mêmes aspects du dépistage ne soient évalués deux fois au travers de plusieurs critères. Les critères et leur description tels qu'ils figurent au Tableau 3 sont une version abrégée des informations par critère. La liste complète originale figure dans le rapport scientifique.


**Tableau 3 – Définition des critères pour l'attribution d'un score par maladie**

Critères	Description
<b>Fréquence (prévalence à la naissance)<sup>c</sup></b>	Prévalence à la naissance en FWB, en VG ou en Belgique si disponible, et/ou en Europe occidentale, en Amérique du Nord et dans le monde
<b>Gravité de la maladie en l'absence de traitement</b>	Gravité de la maladie en l'absence de traitement précoce (évolution naturelle), y compris les séquelles, la mortalité et la perte de qualité de vie qui en découlent
<b>Obtention des résultats du test en temps utile</b>	Les résultats du test sont-ils disponibles suffisamment rapidement pour prévenir des complications évitables et des séquelles ?
<b>Efficacité d'un traitement précoce vs. tardif, suivant le type de traitement (spécifique ou non spécifique)</b>	Quelle est l'efficacité d'un traitement précoce en comparaison avec un traitement tardif ? Existe-t-il des preuves que les participants (nouveau-nés) retireront probablement un bénéfice de ce dépistage et que celui-ci ne leur causera aucun dommage, par exemple au travers d'un diagnostic plus précoce mais sans impact significatif sur l'évolution clinique? Degré de consensus autour du circuit diagnostique à suivre en cas de test positif. Degré de consensus autour de la prise en charge de la maladie si le diagnostic est confirmé. Disponibilité de structures de diagnostic et de traitement.
<b>Probabilité et impact des résultats faux-positifs</b>	Des résultats faux-positifs ont-ils été signalés, à quelle fréquence et quel en est l'impact ?
<b>Probabilité et impact des résultats faux-négatifs</b>	Des résultats faux-négatifs ont-ils été signalés, à quelle fréquence et quel en est l'impact ?
<b>Impact sur le système de soins de santé</b>	Impact sur la capacité diagnostique et thérapeutique, y compris la prise en charge des (vrais) cas dépistés et des diagnostics fortuits. Coût (de l'ajout) d'une maladie spécifique dans le programme existant. Coût diagnostique des tests de confirmation en cas de dépistage positif (vrais et faux positifs). Coût de la prise en charge des cas confirmés (pour les payeurs et pour la société). Rapport coût-efficacité du dépistage éventuel de cette maladie (si disponible). Communication vis-à-vis des participants (parents) concernant les bénéfices et dommages potentiels associés au dépistage d'une maladie supplémentaire. Aspects organisationnels du dépistage et faisabilité des examens ultérieurs. Acceptabilité de la stratégie diagnostique pour la population.

<sup>c</sup> Prévalence à la naissance: voir le rapport scientifique pour la définition



Pour deux critères quantitatifs (la fréquence de la maladie et la disponibilité des résultats en temps utile), une fourchette de valeurs a été définie au préalable par le groupe de pilotage pour favoriser une certaine homogénéité des scores parmi les évaluateurs (Tableau 4).

**Tableau 4 – Recommandations pour l’attribution d’un score pour deux critères quantitatifs**

Score	Fréquence de la maladie : Prévalence à la naissance (pour 100 000 naissances vivantes)	Disponibilité des résultats du test en temps utile : pourcentage de résultats disponibles suffisamment rapidement pour prévenir des complications graves
1	<1	< 50%
2	1 à 5	50% à <80%
3	>5 à 10	80% à <100%
4	>10/100 000	100%

## 4. DÉTERMINATION DU POIDS DE CHAQUE CRITÈRE

À l’aide de l’échelle de LIKERT en quatre points, chaque membre du groupe de pilotage a attribué à chaque critère une pondération reflétant son importance pour la prise de décision. Les résultats préliminaires ont ensuite été discutés au sein du groupe de pilotage puis pondérés une seconde fois après clarification et définition plus précise d’un certain nombre de critères.

Les poids moyen et médian attribués à chaque critère par les évaluateurs ne comportent pas de grandes différences (voir Figure 1), malgré une tendance à accorder une pondération un peu plus importante aux critères « gravité de la maladie », « efficacité du traitement précoce vs tardif » et « probabilité et impact des faux-négatifs ». Ces mesures globales cachent toutefois l’hétérogénéité des pondérations données par chaque évaluateur (voir Figure 2) : si le critère « gravité de la maladie » était unanimement considéré comme très important, les réponses concernant « la fréquence de la maladie », « la disponibilité des résultats en temps utile » et « la probabilité et l’impact des faux-négatifs » étaient beaucoup plus divergentes. Ces différences entre pondérations ont été débattues au sein du groupe de pilotage. Elles peuvent être dues aux facteurs suivants :

- Bien que chaque critère ait été défini et détaillé, des malentendus ont subsisté sur leur contenu précis. Les membres du groupe de pilotage ont proposé de donner désormais une description encore plus précise du contenu de chaque critère (p.ex. sous forme de check-list) et de l’expliquer oralement aux évaluateurs.
- Il est probable que le profil des évaluateurs ait influencé la pondération attribuée à chaque critère : des preneurs de décision, des experts des laboratoires de DNS, des experts cliniques, des représentants des patients et des experts en éthique n’ont pas nécessairement tous les mêmes priorités. Le critère « disponibilité des résultats en temps utile », par exemple, était moins important pour certains experts de laboratoire, parce qu’ils adaptent automatiquement à chaque maladie spécifique la vitesse à laquelle ils communiquent les résultats (p.ex. par téléphone plutôt que par la poste). Lors d’exercices futurs de définition des priorités, il faudra être attentif à une répartition équilibrée de ces profils dans le groupe d’évaluation.



- Le critère « impact sur le système de soins de santé » a une signification très vaste englobant des questions à la fois économiques, organisationnelles, de faisabilité et d'acceptabilité, ce qui le rend difficile à pondérer (et à quantifier pour chaque maladie). De plus, il inclut à la fois le dépistage et la prise en charge des cas identifiés, alors que ces aspects sont organisés de façon différente à ces deux niveaux du système de santé : les politiques de dépistage font partie de la prévention en santé et donc des compétences des Communautés, tandis que la prise en charge des cas relève de la médecine curative et donc de la sécurité sociale fédérale (INAMI).



Figure 1 – Poids (moyen et médian) des sept critères (sur une échelle de LIKERT de 1 à 4 points)

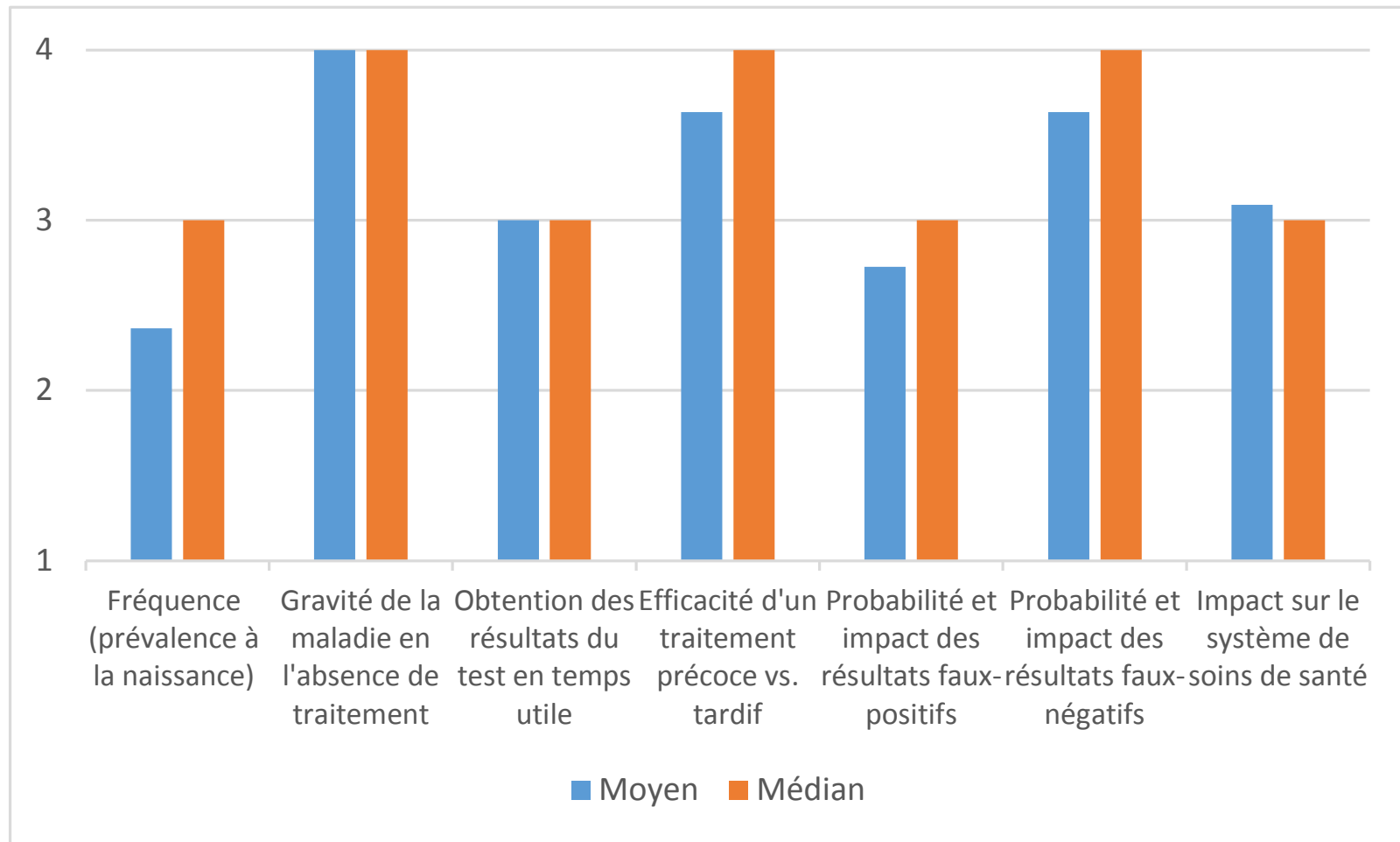
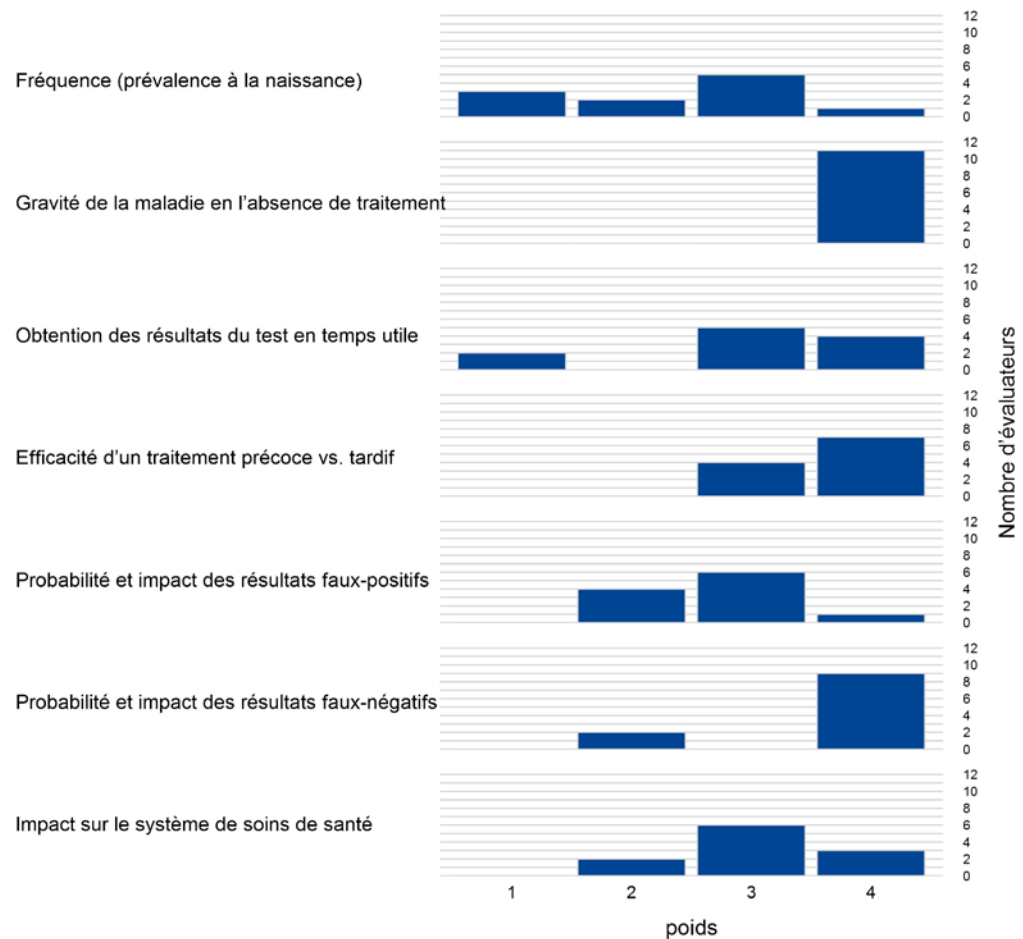




Figure 2 – Distribution des poids attribués aux critères par les différents évaluateurs.





## 5. COLLECTE D'INFORMATIONS SUR CHAQUE PATHOLOGIE

Pour les six maladies sélectionnées, le KCE a mené une étude pragmatique de la littérature existante pour collecter une série d'informations relatives à chaque critère : données belges issues des Communautés, informations en provenance d'autres agences (internationales) et des évaluateurs du groupe de pilotage, littérature scientifique et littérature grise. Étant donné la faible fréquence de ces pathologies, les données disponibles n'étaient pas toujours suffisantes ; en particulier, les données concernant l'impact sur le système de soins étaient assez maigres. L'information a ensuite été résumée dans une courte fiche synthétique sur chaque maladie et fournie aux évaluateurs pour servir de base commune et objective à l'attribution des scores.

## 6. ATTRIBUTION DES SCORES PAR PATHOLOGIE

Les évaluateurs ont attribué aux six maladies un score pour chacun des sept critères, soit un total de 42 scores par évaluateur. Les résultats préliminaires ont été soumis pour validation aux membres du groupe de pilotage. Les informations sur les critères et maladies ont été re-clarifiées et les évaluateurs ont eu la possibilité d'adapter leurs scores en cas de malentendus.

La Figure 3 reprend les scores moyens et médians par critère (sans la pondération par critère) pour chacune des maladies. On observe un certain nombre de tendances, avec des scores généralement élevés (médiane  $\geq 3$ ) pour les critères « gravité de la maladie », « disponibilité des résultats en temps utile » et « efficacité d'un traitement précoce vs. tardif » et au contraire plus faibles (médiane de 1-2) pour la probabilité et l'impact des faux-négatifs et faux-positifs (sauf pour la tyrosinémie de Type I).

La distribution des scores par critère montre également des différences marquées entre les réponses individuelles, sauf pour le critère « gravité de la maladie » (voir données plus détaillées dans le rapport scientifique). Les différences les plus marquées de score entre les évaluateurs s'observaient pour deux critères, « la probabilité et l'impact des faux-négatifs » et « l'impact sur le système de soins de santé ». Les raisons de ces discordances ont été discutées au sein du groupe de pilotage, ce qui a débouché sur une série d'hypothèses :

- Une première explication possible est que, s'agissant de maladies rares, les informations disponibles soient incomplètes pour certains critères. En l'absence de données suffisamment robustes, les connaissances et l'expérience personnelles des évaluateurs ont probablement eu un impact considérable sur les scores.
- En ce qui concerne la probabilité et l'impact des résultats faux-négatifs, il est possible que la probabilité d'un tel résultat soit élevée alors que son impact est limité. Ces deux aspects peuvent donc difficilement être résumés en un score unique.
- Le critère « impact sur le système de soins de santé » a systématiquement donné lieu à des scores très divergents (de 1 ou 2 à 4), ce qui n'est pas reflété dans les moyennes et médianes des scores,



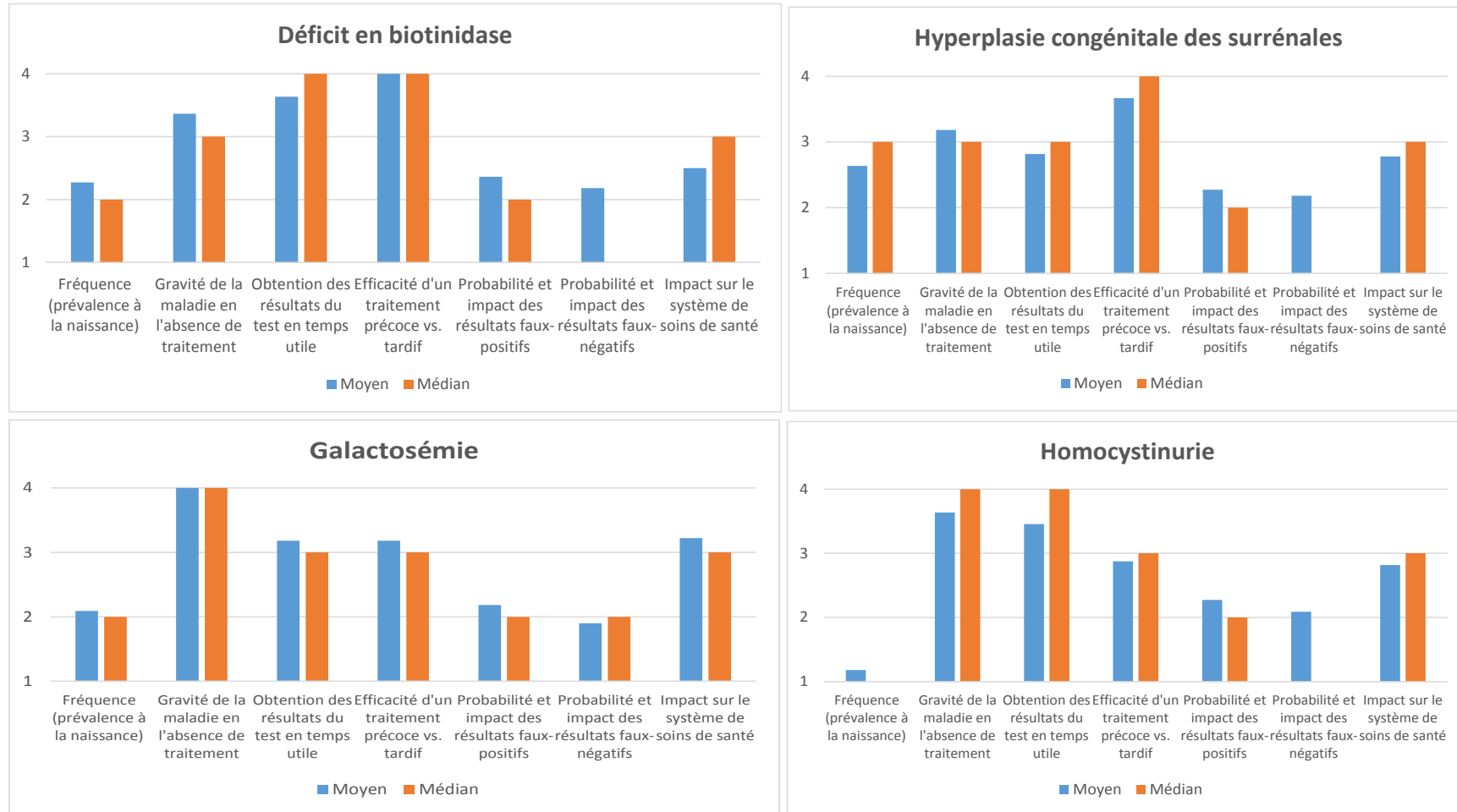


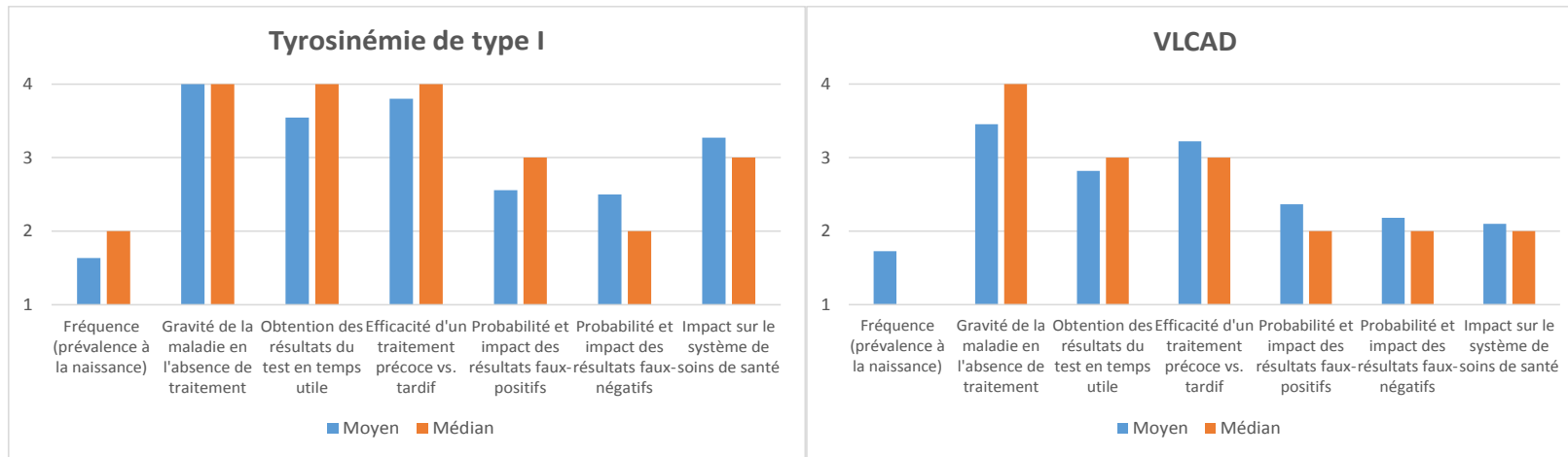
qui tournent généralement autour de 3. Ceci s'explique par la définition très large de ce critère, qui recouvre des composantes extrêmement diverses (Tableau 3), mais aussi par les différents profils des évaluateurs. Ceux qui sont impliqués dans la prise de décision étaient par exemple plus sensibles aux questions économiques et budgétaires que les experts de laboratoire, qui s'intéressaient davantage aux caractéristiques du test de dépistage.

- Il est possible que les représentants des Communautés et des laboratoires DNS accordent plus d'importance à l'impact sur la détection précoce, tandis que d'autres acteurs seraient plus sensibles à l'impact sur les soins de santé curatifs.



Figure 3 – Scores moyens et médians par maladie pour chacun des critères (sans pondération pour l'importance de ces derniers)





VLCAD: Déficit en acyl-CoA déshydrogénase des acides gras à très longue chaîne



## 7. SCORE GLOBAL PAR MALADIE

Le constat le plus marquant de cette étude pilote d'analyse décisionnelle multicritère est que les scores globaux diffèrent très peu par maladie dans l'analyse de base, ce qui rend difficile un classement par ordre de priorité (Figure 4). Les scores médians globaux vont de 2,9 à 3,1 avec une grande dispersion entre les scores individuels (calculés par évaluateur).

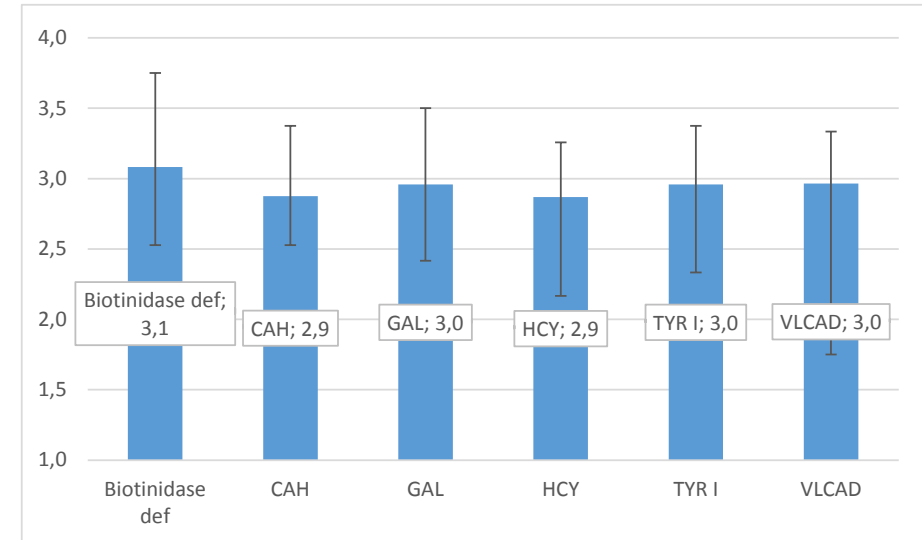
Comme nous l'avons expliqué dans la section sur la méthodologie, nous avons également exploré cinq autres méthodes de calcul des scores globaux par pathologie afin de pouvoir les comparer pour chacune d'elles (voir la section Champ et méthodologie). La Figure 5 révèle que le classement de chaque maladie dépend en partie de la méthode de calcul des scores globaux. Nous avons néanmoins pu dégager certaines tendances ; le déficit en biotinidase est classé en première ou deuxième position quelle que soit la méthode utilisée, la GAL arrive en première ou seconde place avec trois des six approches et le HCY et VLCAD figurent presque systématiquement dans le bas du classement.<sup>d</sup>

Les raisons qui peuvent expliquer ces faibles différences entre les scores globaux ont également été discutées au sein du groupe de pilotage :

- Les six maladies présentent probablement une importance et un intérêt similaires pour le dépistage DNS dans notre pays ; dans la mesure où elles ont déjà été sélectionnées pour le dépistage dans l'une des Communautés, on peut en effet conclure qu'elles ont été jugées suffisamment importantes.
- Il est possible que la fourchette de scores relativement étroite (de 1 à 4) diminue les différences potentielles entre scores individuels. Une échelle de 1 à 10 pourrait éventuellement produire des résultats différents.

- Les indicateurs utilisés, la médiane ou la moyenne (« mesures de tendance centrale ») peuvent masquer certaines différences entre les scores individuels.

**Figure 4 – Analyse de base : scores globaux par pathologie (et intervalle entre les scores individuels des évaluateurs)**

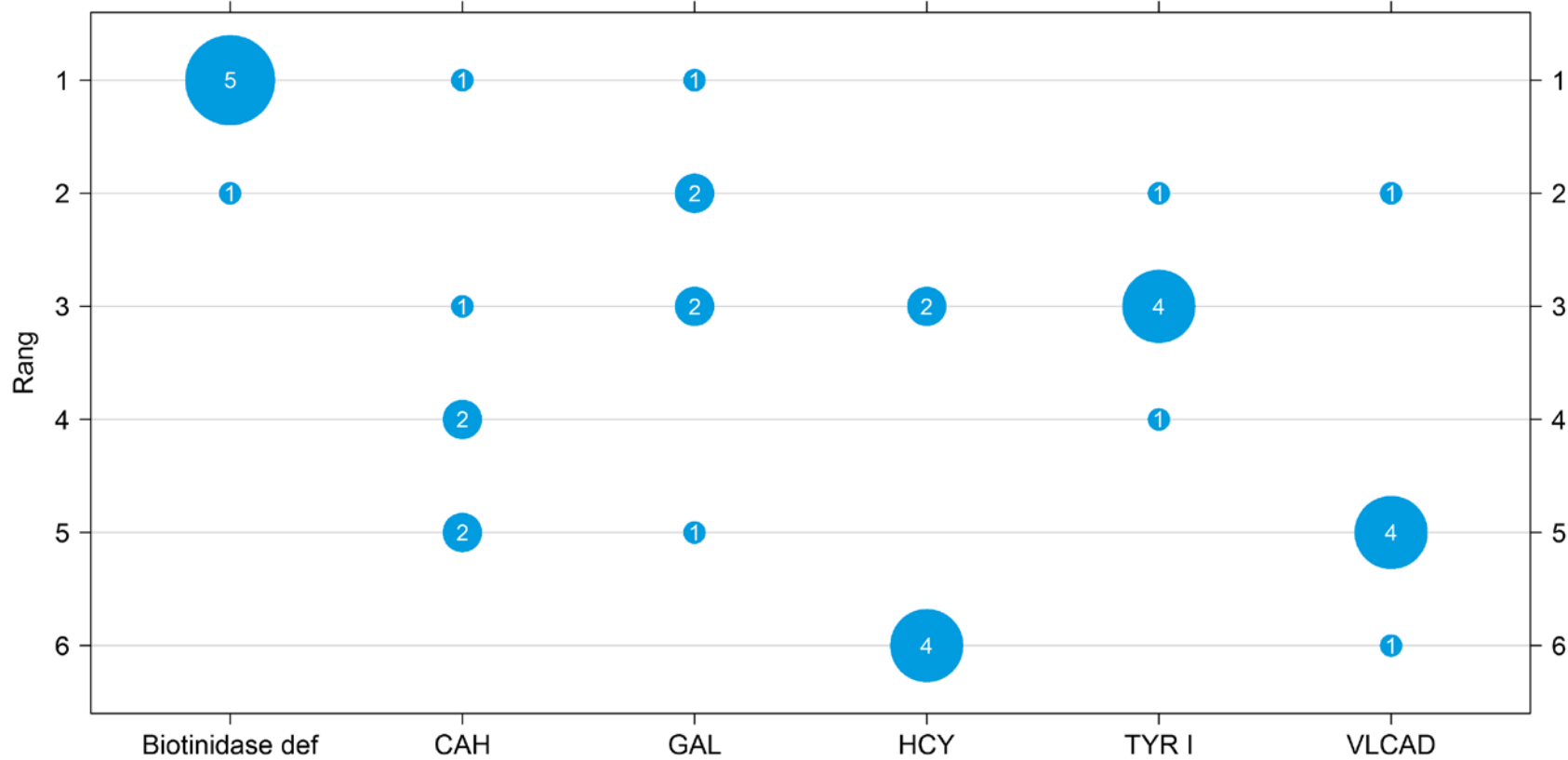


*Biotinidase def.* : Déficit en biotinidase ; *CAH* : Hyperplasie congénitale des surrénales ; *GAL* : Galactosémie ; *HCY* : Homocystinurie ; *TYR I* : Tyrosinémie de type I ; *VLCAD* : Déficit en acyl-CoA déshydrogénase des acides gras à très longue chaîne

<sup>d</sup> En raison des ex-aequo, les totaux ne sont pas toujours six par classement



Figure 5 – Analyse complémentaire : répartition du classement par pathologie pour les six différentes méthodes de calcul des scores globaux



*Biotinidase def. : Déficit en biotinidase ; CAH : Hyperplasie congénitale des surrénales ; GAL : Galactosémie ; HCY : Homocystinurie ; TYR I : Tyrosinémie de type I ; VLCAD : Déficit en acyl-CoA déshydrogénase des acides gras à très longue chaîne*



## 8. CONCLUSION

Nous avons testé dans cette étude une méthode d'analyse décisionnelle multicritère pour l'attribution d'un ordre de priorité à un certain nombre de maladies détectables par dépistage néonatal sanguin. Dégager un consensus sur la sélection et la définition uniforme des critères s'est avéré plus complexe que prévu et a imposé un investissement conséquent en temps et en concertation. Les différences observées entre les scores globaux des différentes maladies dans le cadre de cette étude étaient extrêmement faibles, et donc leur pouvoir discriminant pour servir de base à un classement dans l'optique du DNS est également faible.

Le classement diffère aussi légèrement suivant la méthode de calcul des scores globaux, ce qui souligne le besoin de recherches et discussions supplémentaires sur le sujet.

Plusieurs membres du groupe de pilotage ont également fait remarquer que, s'agissant d'une matière aussi complexe, il n'était pas approprié de baser les politiques et la prise de décision uniquement sur un score global par maladie. Dans la mesure où la pondération des critères et l'attribution des scores encourage la réflexion et la formulation d'arguments explicites, le groupe de pilotage a néanmoins jugé que les résultats de cet exercice ont leur utilité comme aide à la prise de décision.

La méthode utilisée pourrait être améliorée par une définition plus précise de chaque critère, par une division des critères qui comportent des composantes antagonistes en plusieurs critères, par l'application éventuelle d'une échelle plus large pour la pondération et l'attribution des scores et par une prise en compte prudente de l'équilibre entre les profils des évaluateurs.

Les études futures de classement des maladies devraient idéalement porter sur un plus grand nombre de pathologies et pourront s'appuyer sur les enseignements de cette première étude pilote.

## 9. MESSAGES-CLÉS ET ENSEIGNEMENTS POUR L'AVENIR

- La prise de décision quant à l'ajout éventuel d'une maladie (supplémentaire) au programme de dépistage néonatal est un processus complexe qui recouvre plusieurs dimensions. De tels processus décisionnels complexes s'appuient de plus en plus souvent sur des méthodes d'évaluation formalisées.
- En nous basant sur l'expérience de l'INESSS au Québec, nous avons réalisé une étude pilote pour évaluer l'utilité et la faisabilité d'une analyse décisionnelle multicritère (MCDA). Ensemble avec un groupe de pilotage, composé de représentants et d'experts des Communautés et d'autres intervenants, nous avons décidé de tester la méthode sur les six maladies qui ne sont actuellement dépistées que dans l'une des deux Communautés.
- La sélection des critères d'évaluation est loin d'être simple. Les scores parfois très hétérogènes obtenus pour la pondération de certains critères indiquent qu'il vaut mieux, dans le cadre de ce processus, effectuer plusieurs tours successifs en précisant les critères, voire en les complétant ou en les divisant si nécessaire.
- Les évaluateurs adoptent parfois des positions extrêmement différentes. L'attribution d'un poids et d'un score à chacun des critères retenus demande un vaste débat.
- L'attribution d'un score à chaque critère pour chacune des maladies impose de rechercher au préalable les données probantes disponibles sur leur épidémiologie, les tests disponibles et les interventions existantes, en ce compris les aspects touchant à l'organisation, à l'économie de la santé et à l'éthique. Ces informations doivent être communiquées à tous les évaluateurs afin qu'ils puissent tous établir leurs scores sur les mêmes données de base.
- Une échelle de LIKERT à quatre points pour pondérer les critères et attribuer des scores à chaque maladie n'autorise peut-être pas une discrimination suffisamment précise. Nous n'avons toutefois pas pu trancher cette question avec certitude. En effet, sachant que les six maladies étudiées ont été sélectionnées par l'une des Communautés mais pas par l'autre, elles se trouvaient probablement toutes dans la



même zone entre ce qui est justifié ou non dans un programme de dépistage.

- Tous les points de vue pertinents doivent être représentés dans la composition du panel d'évaluateurs, afin de ne passer à côté d'aucun argument essentiel.
- Les discussions dans le groupe de pilotage suggèrent qu'il serait peut-être utile de prévoir dans l'échelle d'évaluation des maladies un « score d'exclusion » indiquant qu'une condition nécessaire à son dépistage n'est pas remplie.
- Le mode de calcul des scores globaux par maladie peut influencer leur classement définitif. Ici non plus, les six pathologies retenues n'ont toutefois pas permis de tirer des conclusions plus précises.
- L'intention n'est pas que la décision d'inclure ou d'exclure certaines maladies du programme de dépistage se prenne uniquement sur la base d'un classement MCDA. Un tel exercice permet toutefois un débat plus objectif et plus transparent autour de la prise de décision, et donc une plus grande cohérence entre des décisions successives.



## ■ RECOMMANDATIONS<sup>e</sup>

### *Aux autorités compétentes des entités fédérées*

- Il n'existe pas de consensus international sur les maladies qui devraient être incluses dans un programme de dépistage néonatal sanguin ni sur comment ces décisions devraient être prises. Pour ces questions complexes, une méthode d'analyse de décision multi-critère (MCDA) peut servir d'outil supplémentaire dans la prise de décision.
- Le MCDA ne peut pas être utilisé comme un outil de décision mathématique en soi, mais cette méthode peut aider à structurer le processus de décision et à le rendre plus transparent. L'application systématique d'une telle méthode peut aussi, à terme, mener à une plus grande cohérence dans la prise de décision.
- Les critères utilisés doivent au préalable être discutés et définis de manière approfondie avec les décideurs et l'ensemble des évaluateurs. Certains critères qui sont liés aussi bien à la gravité de la maladie qu'au traitement possible pourraient être scindés en plusieurs éléments car ceux-ci peuvent être antagonistes.
- La meilleure méthode de calcul des scores globaux pour ce type d'analyse MCDA doit être davantage explorée.
- Pour chaque maladie évaluée, il est nécessaire de rassembler l'information utile sur chaque critère et de la mettre à disposition des évaluateurs.
- Lors de la constitution du groupe d'évaluateurs, il faut veiller à l'équilibre entre les différentes parties impliquées (décideurs, experts, représentants des patients, éthiciens etc.) pour que toutes les opinions et préférences soient le mieux possible représentées.
- Il est difficile d'obtenir de l'information sur les coûts et les bénéfices du trajet diagnostique après un résultat de dépistage positif (y compris les faux positifs) et du traitement ultérieur des vrais positifs. Nous recommandons de prévoir un suivi systématique et à long terme de tous les nouveau-nés qui ont été testés positifs pour pouvoir mieux y répondre dans le futur.

---

<sup>e</sup> Le KCE reste seul responsable des recommandations.







## COLOPHON

Titre :	Dépistage néonatal sanguin : analyse de décision multicritère pour sélectionner les maladies prioritaires – Synthèse
Auteurs :	Chris De Laet (KCE), Germaine Hanquet (KCE), Erik Hendrickx (KCE)
Coordinateur de projet :	Marijke Eyssen (KCE)
Relecture :	Irina Cleemput (KCE), Frank Hulstaert (KCE), Raf Mertens (KCE), Genevieve Veereman (KCE)
Experts externes et stakeholders:	François Boemer (Université de Liège), Marc Boënné (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), Linda De Meirleir (AZ-VUB), François Eyskens (Provinciaal Centrum voor de opsporing van metabole aandoeningen), Philippe Goyens (CHU Brugmann et HUDERF), Hilde Laeremans (AZ-VUB et 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) et 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 et ONE).
Validateurs externes :	Jean-Jacques Cassiman (KU Leuven), Corinne De Laet (HUDERF), Brigitte Côté (Institut National d'Excellence en Santé et en Services Sociaux (Québec)).
Remerciements :	Muriel Deguerry (Observatoire de la Santé et du Social de Bruxelles-Capitale) pour les informations sur le dépistage néonatal en région de Bruxelles Capitale. Stephan Devriese (KCE) pour son aide dans la préparation de certains graphiques. Karin Rondia (KCE) et Gudrun Briat (KCE) pour la révision de la synthèse en français et néerlandais. Photos de couverture : De Kern – verloskundigenpraktijk (NL) et Andrée Mangin (KCE).
Autres intérêts déclarés :	Aucun intérêt déclaré : 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). Appartenance à un groupe de parties prenantes pour lequel les résultats de ce rapport pourraient avoir un impact : Hilde Laeremans (ULB et Vlaams Centrum Brussel voor opsporing van aangeboren Metabole Aandoeningen), Tatiana Pereira (FWB, Direction Générale de la Santé Exécutant pour ONE) and Roland Schoos (Université de Liège, Centre de dépistage néonatal de Liège). Présidence ou fonction de responsable au sein d'une institution, d'une association, d'un département ou d'une autre entité pour lequel/laquelle les résultats de ce rapport pourraient avoir un impact : Karen Colaert (organise le programme de dépistage en Flandre "aangeboren aandoeningen bij pasgeborenen via een bloedstaal" pour la



Vlaams Agentschap Zorg en Gezondheid), Lutgarde De Baere (présidente de BOKS, une association de patients pour les maladies métaboliques), Philippe Goyens et Hilde Laeremans (coordination d'un laboratoire de dépistage des maladies métaboliques pour l'ULB et l'AZ-VUB (VCBMA)), Roland Schoos (participe au comité de pilotage pour le dépistage néonatal de la FWB), Béatrice Toussaint (responsable du dépistage néonatal au sein du cabinet de la FWB en 2009-2014), Pieter Vandenbulcke (responsable de l'équipe prévention générale à la Vlaams Agentschap Zorg en Gezondheid, entre autres pour les programmes de dépistage de la population liés à la prévention des maladies).

Autres intérêts possibles qui pourraient mener à un conflit d'intérêts potentiel ou réel : Brigitte Côté (participation comme investigateur principal à l'étude scientifique de l'INESSS (2013) "Pertinence d'élargir le programme de dépistage néonatal sanguin au Québec")

Layout :

Joyce Grijseels, Sophie Vaes.

**Disclaimer :**

- **Les experts externes ont été consultés sur une version (préliminaire) du rapport scientifique. Leurs remarques ont été discutées au cours des réunions. Ils ne sont pas co-auteurs du rapport scientifique et n'étaient pas nécessairement d'accord avec son contenu.**
- **Une version (finale) a ensuite été soumise aux validateurs. La validation du rapport résulte d'un consensus ou d'un vote majoritaire entre les validateurs. Les validateurs ne sont pas co-auteurs du rapport scientifique et ils n'étaient pas nécessairement tous les trois d'accord avec son contenu.**
- **Enfin, ce rapport a été approuvé à l'unanimité par le Conseil d'administration (voir <http://kce.fgov.be/fr/content/le-conseil-dadministration-du-centre-dexpertise>).**
- **Le KCE reste seul responsable des erreurs ou omissions qui pourraient subsister de même que des recommandations faites aux autorités publiques.**

Date de publication :

25 avril 2016

Domaine :

Health Services Research (HSR)

MeSH :

Decision Making; Metabolism, Inborn Errors; Neonatal Screening

Classification NLM :

WD 205

Langue :

Français

Format :

Adobe® PDF™ (A4)

Dépot légal :

D/2016/10.273/42

ISSN :

2466-6440



Copyright :

Les rapports KCE sont publiés sous Licence Creative Commons « by/nc/nd »

<http://kce.fgov.be/fr/content/a-propos-du-copyright-des-rapports-kce>



Comment citer ce rapport ?

De Laet C, Hanquet G, Hendrickx E. Dépistage néonatal sanguin : analyse de décision multicritère pour sélectionner les maladies prioritaires – Synthèse. Health Services Research (HSR). Bruxelles: Centre Fédéral d'Expertise des Soins de Santé (KCE). 2016. KCE Reports 267Bs. D/2016/10.273/42.

Ce document est disponible en téléchargement sur le site Web du Centre Fédéral d'Expertise des Soins de Santé.