



HAL
open science

**Les mouvements généraux : Constitution et validation
d'une mesure neuromotrice pour développer un système
d'évaluation automatisée destinée au nouveau-né
prématuré**

Angelica Valencia Diaz

► **To cite this version:**

Angelica Valencia Diaz. Les mouvements généraux : Constitution et validation d'une mesure neuromotrice pour développer un système d'évaluation automatisée destinée au nouveau-né prématuré. Psychologie. Université Paris sciences et lettres, 2021. Français. NNT : 2021UPSLP022 . tel-03566090

HAL Id: tel-03566090

<https://theses.hal.science/tel-03566090>

Submitted on 11 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'École Pratique des Hautes Études

**Les mouvements généraux : constitution et validation
d'une mesure neuromotrice pour développer un système
d'évaluation automatisée destinée au nouveau-né
prématuré**

Soutenue par

Angélica VALENCIA-DÍAZ

Le 16 03 2021

École doctorale n° 472

**École doctorale de l'École
Pratique des Hautes Études**

Spécialité

Psychologie cognitive



École Pratique
des Hautes Études

PSL 

Composition du jury :

Elsa, KERMORVANT, Professeur HDR, Université de Paris et AP-HP	<i>Présidente du Jury</i>
Jean, BERGOUGIOUX, Professeur PHPH AP-HP Raymond-Poincaré	<i>Rapporteur</i>
Véronique, PIERRAT MD, PhD Chef de Service Néonatalogie CHU Lille - INSERM-U 1153	<i>Rapporteur</i>
Sylvie, VIAUX-SAVELON, Pédiopsychiatre HDR, APHP, Pitié Salpêtrière-Trousseau	<i>Examineur</i>
Emmanuel, DEVOUCHE Maître de Conférences HDR, Université de Paris	<i>Examineur</i>
Delphine, MITANCHEZ, Professeur Chef de service Néonatalogie, CHRU Tours	<i>Examineur</i>
Joëlle, PROVASI Maître de Conférences HDR, EPHE-PSL	<i>Co-Directrice thèse</i>
François, JOUEN Doyen de la section SVT, EPHE-PS	<i>Directeur de thèse</i>

Résumé

Le nouveau-né prématuré présente un risque neurologique dont la gravité augmente avec des complications cliniques. L'estimation du risque nécessite des outils d'évaluation neuromotrice qui orientent un suivi précoce et adapté aux contraintes de la prématurité. L'analyse des mouvements généraux (GMs) est l'outil le plus adapté, car il reflète l'intégrité du système nerveux sans manipuler l'enfant. Cependant, deux problèmes ne favorisent pas l'utilisation de cet outil : la forte implication de la subjectivité de l'évaluateur et le manque d'études de validation pour le nouveau-né prématuré. Cette thèse présente deux études sur la constitution et la validation d'une base de données des GMs qui, en tant que mesure neuromotrice, permettra le développement d'un outil automatique destiné au nouveau-né prématuré. La première étude décrit longitudinalement les GMs au moyen de trois résultats. Premièrement, les GMs ne permettent pas d'identifier de changements développementaux, même si l'implication du cou et tronc, ainsi que la variabilité et la fluidité sont plus importantes à la période terme qu'à la période préterme. Deuxièmement, les GMs tracent l'évolution clinique, car le taux des GMs atypiques diminue aux périodes terme et post terme. Troisièmement, les GMs permettent de distinguer les enfants ayant des anomalies motrices et ceux n'en ayant pas. Ces résultats doivent s'interpréter avec précaution à cause d'une perte de données dans le suivi longitudinal. La deuxième étude teste la validité et fiabilité des GMs au moyen de trois résultats. Premièrement, la validité du construit des GMs fut confirmée par des associations avec une échographie cérébrale atypique et l'hémorragie intraventriculaire. Deuxièmement, les GMs présentent une bonne sensibilité (0.87 – 0.90) pour identifier les enfants ayant une lésion cérébrale, mais une mauvaise spécificité (0.28 – 0.31) pour identifier ceux n'en ayant pas. Troisièmement, la fiabilité intra/inter codeur varie entre modérée et bonne. La généralisation de ces résultats doit se faire avec précaution, car la sensibilité et la spécificité dépendent des caractéristiques de l'échantillon.

Mots clés : Nouveau-né prématuré, Mouvements généraux (GMs), Mesure neuromotrice, Évaluation des GMs automatisée, Validité, Fiabilité

Abstract

The premature newborn presents a neurological risk that increases in severity with clinical complications. Risk assessment requires neuromotor assessment tools that guide early follow-up adapted to the constraints of prematurity. General Movement Analysis (GMs) is the most suitable tool, as it reflects the integrity of the nervous system without manipulating the child. However, two problems do not favour the use of this tool: the strong involvement of the evaluator's subjectivity and the lack of validation studies for the premature newborn. This thesis presents two studies on the constitution and validation of a database of GMs which, as a neuromotor measure, will allow the development of an automatic tool for the premature newborn. The first study describes GMs longitudinally using three results. Firstly, the GMs do not allow the identification of developmental changes, even though the involvement of the neck and trunk, as well as variability and fluidity are more important in the term period than in the preterm period. Second, GMs trace clinical evolution, as the rate of atypical GMs decreases in the term and post-term periods. Thirdly, GMs help distinguish between children with and without motor abnormalities. These results should be interpreted with caution due to high attrition. The second study tests the validity and reliability of the GMs using three outcomes. First, the validity of the GM construct was confirmed by associations with atypical brain ultrasound and intraventricular hemorrhage. Second, the GMs have good sensitivity (0.87 - 0.90) for identifying children with brain injury, but poor specificity (0.28 - 0.31) for identifying those without brain injury. Third, intra/inter coder reliability varies between moderate and good. Care should be taken in generalizing these results, as sensitivity and specificity depend on the characteristics of the sample.

Key Words : Premature Newborn, General Movement (GMs), Neuromotor Assessment, Automated GMs Assessment, Validity, Reliability.

Remerciements

À l'EPHE, d'avoir financé cette thèse au moyen d'un contrat doctoral.

À M. François JOUEN, de m'avoir proposé ce beau sujet de thèse.

À Me. Joëlle PROVASI, pour sa constante disponibilité et précieuses suggestions.

Aux membres du jury qui ont accepté d'évaluer ce travail de thèse.

À Pr. Delphine MITANCHEZ, pour son accueil bienveillant au service de néonatalogie à l'hôpital Armand Trousseau.

À Dr. Dominique BOHU, pour sa grande générosité et expertise lors des observations à l'hôpital Armand Trousseau.

À l'équipe médicale, soignante et administrative de l'hôpital Armand Trousseau, pour sa collaboration lors du recrutement et collecte des données.

À Pr. Elsa KERMORVANT, pour son accueil au service de néonatalogie à l'hôpital Necker Enfants Malades et suggestions toujours pertinentes.

À l'équipe médicale, soignante et administrative de l'hôpital Necker Enfants Malades, pour son professionnalisme et déférence lors des observations.

Aux parents qui ont consenti que leurs enfants participent à cette étude.

À mes ami(e)s et ma famille, pour leur soutien et amour.

À Emmanuel, mon ami, une source infinie.

Sommaire

Introduction	8
I. La naissance prématurée : un enjeu neurologique	11
1.1. Définitions et chiffres	11
1.2. Enjeux neurodéveloppementaux de la prématurité	12
1.2.1. L'extrême prématurité.....	12
1.2.2. La grande prématurité	13
1.2.3. La prématurité moyenne ou tardive	15
1.2.4. Complications cliniques	17
1.2.4.1. Dysplasie bronchopulmonaire	17
1.2.4.2. Entérocolite ulcéro-nécrosante.....	18
1.2.4.3. Sepsis	19
1.2.4.4. Lésions cérébrales.....	21
1.3. Évaluation de la structure cérébrale.....	23
II. L'évaluation neuromotrice	27
2.1. Évaluation de la fonction cérébrale	27
2.2. Les outils d'évaluation neuromotrice	28
2.2.1. Les outils d'évaluation neuromotrice destinés au nouveau-né prématuré	33
2.2.1.1. Le Test of Infant Motor Performance	33
2.2.1.2. L'analyse des mouvements généraux.....	34
III. Les mouvements généraux	38
3.1. Parcours évolutif	38
3.2. Corrélats neuronaux	40
3.2.1. Générateurs centraux de patterns.....	42
3.2.2. Générateurs centraux de patterns chez l'humain	42
3.2.3. Modulation	43
3.2.4. Importance clinique.....	44
3.3. Évaluation automatisée	46
3.3.1. Reconnaissance par détection indirecte	46
3.3.2. Reconnaissance par détection directe	48
3.3.3. Reconnaissance par des systèmes hybrides.....	49

3.3.4. Apprentissage sur des modèles cognitifs et sur des bases de données	50
3.3.5. L'évaluation automatisée destinée au nouveau-né prématuré.....	51
3.3.6. Validité de l'évaluation automatisée	51
IV. Validation des mesures neuromotrices.....	54
4.1. Propriétés psychométriques	54
4.1.1. Validité	55
4.1.2. Fiabilité.....	55
4.2. Indices statistiques de validation.....	56
4.2.1. Indices épidémiologiques	56
4.2.2. Indices d'association.....	57
4.2.3. Indices d'accord	59
4.2.4. Taille d'effet.....	61
4.3. Aspects affectant la validation	63
4.3.1. Échantillonnage	63
4.3.2. Mesure	64
4.3.3. Procédure d'évaluation.....	65
4.3.4. Traitement et rapport des données	65
4.4. Utilité clinique.....	66
V. Bilan théorique, problématique et objectifs.....	68
5.1. Problématique.....	69
5.2. Objectifs	69
VI. Constitution d'une base de données des mouvements généraux.....	71
6.1. Méthodologie	72
6.1.1. Participants	72
6.1.2. Recueil des données	74
6.1.3. Évaluation des mouvements généraux	76
6.1.4. Traitement et analyse des données	77
6.2. Résultats	79
6.3. Discussions.....	92
VII. Validation d'une base de données des mouvements généraux.....	97
7.1. Méthodologie	98
7.1.1. Participants	99
7.1.2. Procédure	99
7.1.3. Traitement et analyse des données	101

7.2. Résultats	103
7.2.1. Validité	103
7.2.2. Fiabilité.....	108
7.3. Discussions.....	112
7.3.1. Validité	113
7.3.2. Fiabilité.....	115
Conclusions.....	118
Références.....	120

Introduction

La naissance prématurée est un problème de santé planétaire qui génère plus de perplexités que de certitudes. Paradoxalement, les avancées médicales ont entraîné une augmentation régulière du taux de naissances prématurées. En effet, environ 15 millions d'enfants naissent avant le terme chaque année, soit 1 sur 10 à une échelle mondiale (Blencowe et al., 2012). Spécifiquement en France, le taux de naissances prématurées a augmenté significativement entre 1995 (4 %) et 2016 (6 %) (Coulm, Bonnet, Blondel, Vanhaesebrouck, Vilain, 2017). Cela signifie que 60 000 bébés naissent prématurément chaque année en France, soit 164 chaque jour. La prématurité reste le principal défi en obstétrique et en néonatalogie, car elle représente la cause principale de mortalité périnatale et de la morbidité neurologique à long terme (McCormick, 1985 ; Blencowe, Vos, et al., 2013). A court terme, la naissance prématurée représente un problème de santé grave, car elle implique des soins intensifs immédiats. Ainsi, en fonction de son état de santé, le nouveau-né prématuré requiert d'une ventilation artificielle, une alimentation par sonde, un contrôle des fonctions respiratoires et cardiaques. De plus, les nouveau-nés prématurés les plus immatures requièrent des soins essentiels et supplémentaires, puisqu'ils tendent à développer des complications de santé. Certes, les complications cliniques constituent la première cause de décès néonatal et parmi les survivants un risque accru de troubles neurologiques (Liu et al., 2012 ; Synnes et al., 2017). À long terme, la prématurité constitue un facteur de risque neurodéveloppemental dans le plan sensoriel, cognitif et moteur qui peut entraîner la nécessité d'aides ou de soins spéciaux à vie (Blencowe, Cousens, et al., 2013). Les troubles cognitifs sont généralement mesurés à l'aide de l'échelle de Bayley qui s'applique jusqu'à 2 ans et des tests d'intelligence, qui vont après cet âge (Natarajan & Shankaran, 2016 ; Patel, 2016). Les troubles moteurs sont mesurés par l'examen neurologique ou des tests standardisés. Les problèmes moteurs englobent des problèmes mineurs et des problèmes plus graves. Parmi les problèmes mineurs se

trouvent les troubles de la coordination motrice qui, sans aucune étiologie pathophysiologique (ex. lésion cérébrale), interfèrent avec la vie quotidienne (Hadders-Algra, 2002). Parmi les troubles graves se trouve la paralysie cérébrale (PC). La PC englobe un groupe de perturbations motrices et posturales causées par une anomalie cérébrale non progressive qui survient dans un cerveau en développement (Palmer, 2004). Le niveau d'un dysfonctionnement moteur est évalué selon le compromis fonctionnel généralement mesuré par le système de classification de la fonction motrice globale (en anglais GMFCS).

Des données montrent qu'une intervention précoce, avant le terme, a un effet bénéfique sur le développement cognitif, indépendamment du type d'intervention pourvue (Blauw-Hospers, de Graaf-Peters, Dirks, Bos, & Hadders-Algra, 2007). En ce qui concerne le développement moteur, une intervention précoce est plus efficace lorsqu'elle commence avant l'âge du terme et vise à imiter l'environnement intra-utérin (Blauw-Hospers & Hadders-Algra, 2005). Par conséquent, l'identification des nouveau-nés prématurés affectés d'un risque accru et qui requièrent d'une intervention précoce est cruciale. Cela repose sur une mesure clinique qui permet la détection des signes neurologiques, et ce dès les étapes très précoces. En effet, la détection des signes neurologiques chez le nouveau-né prématuré se base sur l'observation du comportement neuromoteur. Cependant, l'évaluation neuromotrice chez le nouveau-né prématuré doit relever des défis propres de la prématurité et de la validation de l'évaluation à elle-même. Ces constats impliquent deux aspects concernant l'évaluation neuromotrice chez le nouveau-né prématuré : le manque d'outils adaptés et le manque d'études de validation relatives à l'adéquation des outils existants. De ce fait, l'évaluation neuromotrice du nouveau-né prématuré représente un défi qui doit être relevé par l'adoption d'une approche pluridisciplinaire (Marschik et al., 2017). Par conséquent, ce travail de thèse intègre la première partie d'un travail multidisciplinaire qui associera la clinique et l'informatique pour développer un outil automatisé destiné à l'analyse neuromotrice chez le nouveau-né prématuré. Pour ce faire, cette thèse porte sur la constitution et la validation d'une mesure neuromotrice qui permettra de programmer un système informatisé capable d'identifier les anomalies motrices qui témoignent d'une dysfonction neurologique

chez le nouveau-né prématuré. Plus particulièrement, nous décrivons une mesure neuromotrice et rapportons leurs paramètres psychométriques (validité et fiabilité). Cette mesure neuromotrice consiste en la base de données d'un répertoire moteur global connu comme GMs¹ (mouvements généraux) qui ont été obtenus auprès d'enfants nés prématurément.

¹ De l'anglais « General movements » GMs

I. La naissance prématurée : un enjeu neurologique

1.1. Définitions et chiffres

Selon l'Organisation Mondiale de la Santé (OMS), la prématurité se définit comme toute naissance survenant avant 37 semaines de gestation révolues (soit avant le début du 9^e mois de grossesse). Les semaines de gestation se définissent par les semaines d'aménorrhée (SA) comptées à partir des dernières règles, et indiquent la durée de la gestation in utéro jusqu'à l'accouchement à terme ou prématuré. En ce qui concerne la prématurité, nous parlerons de quatre âges : 1) l'âge gestationnel (AG) exprimé en SA qui indique le terme de la grossesse le jour de la naissance ; 2) l'âge chronologique (AC) qui, exprimé en jours, correspond au délai écoulé depuis la naissance ; 3) l'âge gestationnel corrigé (AGc) qui correspond à l'âge gestationnel plus l'âge chronologique ; 4) l'âge corrigé qui s'exprime en semaines post-terme (SPT) et se calcule à partir de la date théorique de naissance (40 SA). La prématurité recouvre 3 sous-catégories : l'extrême prématurité qui correspond à une naissance survenue avant 28 SA (soit en deçà de 6 mois de grossesse), la grande prématurité correspondant à une naissance entre 28 SA et 32 SA (entre 6 et 7 mois), la prématurité moyenne, voire tardive, qui correspond à une naissance entre 32 SA et 37 SA (entre 7 et 8 mois). Bien que la prématurité se classe en ces 3 catégories, des études concluent que l'âge gestationnel représente plutôt un continuum du plus immature au plus mature (Gill & Boyle, 2017). L'accouchement ne devrait pas avoir lieu avant 39 SA sauf indications médicales. Parmi les causes d'une naissance prématurée figurent les grossesses multiples, le retard dans la croissance intra-

utérine, les infections, le diabète et l'hypertension gestationnelle ; il arrive fréquemment que la cause ne soit pas identifiée. Il y aurait aussi une influence génétique. Le plus souvent, le poids à la naissance est proportionnel à l'âge gestationnel. Pour l'OMS, les enfants à faible poids de naissance sont ceux nés avec un poids inférieur à 2500 g. Les enfants de très faible poids de naissance présentent un poids égal ou inférieur à 1500 g. Les enfants d'un poids extrêmement faible naissent avec un poids moins de 1000 g. Aujourd'hui, la limite de viabilité d'un nouveau-né prématuré (limite en dessous de laquelle le bébé ne sera pas réanimé) est 22 SA — 23 SA (soit en deçà de 5 mois de grossesse) et à un poids de moins de 500 g.

1.2. Enjeux neurodéveloppementaux de la prématurité

La naissance prématurée représente un défi important, car elle comporte des risques de santé à court terme et à long terme. Parmi les conséquences les plus graves qui en découlent figurent les complications cliniques et les séquelles neurodéveloppementales. Dans cette section nous décrirons les principaux facteurs de perturbation neurodéveloppementale.

1.2.1. L'extrême prématurité

La naissance extrêmement prématurée (moins de 28 SA) représente 5 % des accouchements avant le terme, dont la survie s'améliore avec les avancées médicales (Kugelman et al., 2012 ; Stoelhorst et al., 2005). Les survivants à 22 SA sont inhabituels, alors que le taux de survie augmente avec l'âge gestationnel passant de 40 % à 23 SA à ≥ 70 % à 26 SA. Selon une estimation mondiale, plus de 45 % des enfants nés extrêmement prématurés développent des troubles

cognitifs, moteurs et perceptifs (Blencowe, Lee, et al., 2013). Comparés aux enfants nés à terme, ceux nés extrêmement prématurés présentent un risque 3 à 6 fois plus élevé (soit 66 % contre 3 %) (Orchinik et al., 2008 ; Serenius et al., 2016). En effet, le risque et la gravité associés augmentent avec la diminution de l'âge gestationnel (Doyle, Roberts, & Anderson, 2010; Wilson-Costello, Friedman, Minich, Fanaroff, & Hack, 2005). Ainsi, les troubles neuro-développementaux présentent la prévalence suivante : 45 % à 22 SA — 23 SA ; 30 % à 24 SA ; 25 % à 25 SA ; et 20 % à 26 SA (Jarjour, 2015 ; Moore et al., 2013). Plus particulièrement, les troubles moteurs sont légèrement moins fréquents (11 % — 15 %) que les troubles cognitifs (7 % — 12 %) (Mansson & Stjernqvist, 2014). La prévalence de la PC (paralysie cérébrale) varie entre 6.9 % et 3.5 % à 24 SA — 27 SA (Pierrat et al., 2017). Bien que la gravité du risque reste un sujet de débat, car le taux de PC peut refléter un taux de décès encore important, les troubles graves diminuent, alors que les troubles moins graves deviennent plus courants (Nagara et al., 2018 ; Rogers & Hintz, 2016 ; Ballot, Potterton, Chirwa, Cooper, & Hilburn, 2012). À cause de son importante immaturité, l'âge gestationnel s'utilise comme un critère pronostique du développement chez l'enfant né extrêmement prématuré. Cependant, un petit âge gestationnel, à lui seul, ne représente pas un facteur pronostique fiable (Herber-Jonat et al., 2014). Cela signifie que la plupart des outils d'évaluation neurologique, basés sur l'âge gestationnel, présentent des limites importantes pour prédire le devenir neurologique des nouveau-nés extrêmement prématurés (Yeo et al., 2017). C'est ainsi que le besoin d'outils d'évaluation neurologique appropriés pour ce groupe d'enfants reste une question pertinente.

1.2.2. La grande prématurité

La grande prématurité (naissance entre 28 et 32 SA) représente 10 % des accouchements prématurés à une échelle mondiale, dont environ 10 000 surviennent en France (Blencowe, Cousens, et al., 2013 ; Larroque et al., 2004). La

survie des enfants nés grands prématurés varie entre 78 % et 93 %, parmi lesquels le taux de survie est plus bas à moins de 29 SA (entre 35 % et 84 %) (Helenius et al., 2018). Ces variations obéissent à divers facteurs tels que le service hospitalier considéré, les caractéristiques des soins pourvus et les régions concernées (Draper et al., 2009). Selon une estimation régionale et globale, 24 % des grands prématurés développent des déficiences neurodéveloppementales, dont la prévalence diminue avec la diminution de l'âge gestationnel et le poids à la naissance (Blencowe, Vos, et al., 2013 ; Milligan, 2010). Chez les grands prématurés, la sévérité des troubles (cognitifs, moteurs ou sensoriels) se partage en : 25 % légère ; 9 % modérée ; et 5 % grave. Comparés aux enfants nés à terme, les grands prématurés présentent plus de problèmes cognitifs (12 % contre 3 %), alors que la prévalence de PC se situe entre 7 % et 9 % (Fawke, 2007 ; Larroque et al., 2008 ; Pascal et al., 2018). Cependant, des données comparant trois périodes (2000, 2005, 2010) démontrent que la prévalence de PC a considérablement diminué (1 %) (Abily-Donval et al., 2015). Indépendamment de la gravité, 20 % des grands prématurés présentent des anomalies neuromotrices impliquant : perturbations de la motricité fine et globale ; asymétrie des mouvements ; hypotonie du tronc ; hypertonie de l'extenseur du cou (Ferrari et al., 2012 ; Pascal, Govaert, & Broeck, 2018). La grande prématurité est corrélée avec des complications, telles que les lésions cérébrales et les infections, qui entraîneraient des troubles neurodéveloppementaux (Beaino et al., 2010 ; Castaño & Maurer, 2015 ; Lee et al., 2014 ; Parikh, 2016). Malgré ces associations avérées, on ne sait toujours pas lequel des évènements, prénataux ou néonataux, prédit le mieux le devenir neurodéveloppemental (Ranke, Krägeloh-Mann, & Vollmer, 2015). Des études proposent que la prédiction du développement serait mieux établie par des biomarqueurs neuromoteurs : des GMs (mouvements généraux) atypiques, une dystonie transitoire et des troubles posturaux (Ferrari et al., 2012).

1.2.3. La prématurité moyenne ou tardive

Les enfants nés modérément prématurés sont ceux qui naissent à 32 SA — 33 SA, alors que ceux nés tardivement correspondent à plus de 34 SA. Ces deux groupes représentent 84 % des naissances prématurées (12 % et 72 % respectivement) (Natarajan & Shankaran, 2016; Jacob et al., 2017; Williams & Pugh, 2018). Comparés à leurs pairs plus petits, les enfants modérément et tardivement prématurés restent peu étudiés, malgré le fait qu'ils représentent la majorité des naissances prématurées (Cheong et al., 2017). Jusqu'à récemment, les nouveau-nés modérément et tardivement prématurés étaient traités cliniquement comme des enfants nés proches du terme (Morgan & Boyle, 2018 ; Tsai et al., 2012 ; Boyle & Boyle, 2013 ; Machado, Passini Júnior, & Rodrigues Machado Rosa, 2014). Or, ils s'avèrent plus immatures et présentent une mortalité et une morbidité 3 à 6 fois plus fréquentes que les enfants nés à terme (Dong & Yu, 2011 ; Williams & Pugh, 2018). Des déficiences cognitives sont observées entre 5 % et 10 % des enfants nés modérément ou tardivement prématurés (Chan, Leong, Malouf, & Quigley, 2016; Johnson et al., 2015; Shah, Kaciroti, Richards, Oh, & Lumeng, 2016; Woythaler, McCormick, Mao, & Smith, 2015). Cette population présente également un risque de troubles neuromoteurs : 41 % présentent une dysfonction neuromotrice légère (ex. Perturbation de la motricité fine) ; 3 % présentent une perturbation neuromotrice modérée (ex. difficulté dans la coordination motrice) (Arnaud et al., 2007) ; un diagnostic de PC est 3 fois plus observé dans cette population que chez des enfants nés à terme (Cheong et al., 2017 ; Kugelman & Colin, 2013 ; Morgan & Boyle, 2018). Une naissance prématurée, modérée ou tardive, s'associe également avec des anomalies lésionnelles et maturationnelles qui représentent un risque : des lésions de la substance blanche et grise ; une petite taille corticale ; un volume réduit de la substance blanche ; une myélinisation sous-développée ; une immaturité des plis (Mewes et al., 2006 ; Kinney, 2006 ; Walsh, Doyle, Anderson, Lee, & Cheong, 2014). Bien que ces anomalies s'avèrent détectables par l'imagerie par résonance magnétique (IRM), ce n'est pas envisageable d'offrir une évaluation par IRM à tous

les nourrissons nés modérément ou tardivement prématurés étant donné leur grand nombre. Par conséquent, c'est important d'avoir d'autres indicateurs cliniques pour identifier les enfants à risque (Spittle et al., 2017). Ainsi, des études suggèrent les signes suivants : une hypotonie du fléchisseur des membres ; un contrôle de la tête faible ; une qualité réduite des mouvements spontanés ; des étirements rapides et moins généralisés ; un taux important de tremblements et réflexes non optimaux (Romeo et al., 2013). Bien que la pertinence clinique de cet ensemble moteur soit incontestée, il présente deux problèmes. Tout d'abord, ce profil moteur ne correspond pas à tous les âges gestationnels recouverts par la naissance modérément ou tardivement prématurée, car il diffère entre les enfants nés à 34 SA et ceux nés à 35 SA (Romeo et al., 2011). De plus, ce profil moteur reflète un trouble neurologique actuel, mais sa capacité prédictive d'un trouble ultérieur n'est pas toujours évidente (Spittle et al., 2017). Compte tenu de ce qui précède, un outil d'évaluation neuromotrice plus sensible, permettant l'identification de signes précoces, reste un enjeu majeur. Cela se révèle primordial concernant les nouveau-nés modérés ou tardivement prématurés, car : ils constituent la majorité des naissances prématurées ; ils présentent un risque accru de troubles neurodéveloppementaux ; ils représentent une population peu étudiée.

En définitive, la naissance prématurée représente un facteur de risque neurologique, sachant que plus l'enfant naît immature, plus le risque augmente. Cependant, le devenir neurodéveloppemental des enfants nés prématurés se détermine par diverses variables : biologiques (ex. les complications cliniques) ; populationnelles (ex. les centres hospitaliers concernés) ; sociétales (ex. le niveau socioéconomique des parents) ; scientifiques (ex. les avancées médicales). Les avancées médicales entraînent l'augmentation de la survie des enfants nés prématurément. Cependant, cela implique l'augmentation du nombre d'enfants concernés par des troubles développementaux. Par ailleurs, bien qu'un âge gestationnel petit et des signes moteurs indiqueraient un risque, ces facteurs restent peu sensibles pour prédire le développement. C'est pourquoi l'identification des nouveau-nés prématurés les plus vulnérables représente un défi à relever.

1.2.4. Complications cliniques

Étant donné qu'une naissance avant le terme interrompt le développement *in utero*, les organes du nouveau-né prématuré sont présents, mais encore immatures. Ceci concerne principalement les poumons, le tube digestif, le cœur et le cerveau. Parmi d'autres facteurs, l'immaturation de ces organes et du système immunitaire peut entraîner les complications suivantes : la dysplasie bronchopulmonaire (DBP), l'entérocolite ulcéro-nécrosante (EUNC), les infections, particulièrement le sepsis, et les lésions cérébrales. Nous allons par la suite décrire chacune de ces complications et leur impact sur le neurodéveloppement du nouveau-né prématuré.

1.2.4.1. Dysplasie bronchopulmonaire

La DBP (dysplasie bronchopulmonaire) est la pathologie la plus fréquemment associée à la prématurité, dont l'incidence augmenta avec la survie accrue d'une naissance extrêmement prématurée (Bhandari & Panitch, 2018 ; Simpson, Hall, & Wilson, 2015 ; Stoll et al., 2015). Cela signifie que la prévalence de la DBP augmente avec la diminution de l'âge gestationnel et du poids à la naissance : Il est de 50 % chez les nouveau-nés avec un poids extrêmement faible pour passer à 5 % chez ceux ayant un poids très faible (Eber & Zach, 2001). La DBP est une altération alvéolaire et vasculaire des poumons, dont l'origine réside dans une complication associée aux soins pulmonaires pour traiter une détresse respiratoire (Berkelhamer, Mestan, & Steinhorn, 2018). En général, la DBP obéit à deux facteurs : à l'interruption du développement normal des poumons ; aux soins médicaux pourvus au nouveau-né prématuré malade (Simpson et al., 2015). Par conséquent, le diagnostic et la gravité de la DBP s'établissent en considérant deux gestes médicaux : l'apport d'oxygène supplémentaire et la ventilation artificielle (Day & Ryan, 2017; Hines et al., 2017; Jobe & Bancalari, 2001; Jobe & Steinhorn, 2017).

Les nouveau-nés prématurés avec DBP sont susceptibles de rencontrer d'autres difficultés graves, comme les lésions cérébrales (Chiriboga et al., 2010 ; Garg, Kurzner, Bautista, & Keens, 1988 ; Skidmore, Rivers, & Hack, 1990). La DBP représente un facteur néonatal qui fait augmenter le risque de troubles neurodéveloppementaux à long terme (Ehrenkranz et al., 2005; Kraft, Verhage, den Heijer, & Bos, 2019; Majnemer et al., 2000). Ainsi, la DBP est donc corrélée à un risque accru de déficiences cognitives telles que les difficultés de la perception, de l'attention, de la mémoire, de l'apprentissage et du langage (Anderson & Doyle, 2006 ; Cheong & Doyle, 2018 ; Potharst et al., 2013 ; Short et al., 2003 ; Singer, Yamashita, Lilien, Collin, & Jill, 1997). De même, la DBP est associée à des troubles de la motricité fine et globale (Bos, Van Braeckel, Hitzert, Tanis, & Roze, 2013 ; Bose et al., 2009 ; Ehrenkranz et al., 2005 ; Hintz, Kendrick, Vohr, & Poole, 2005 ; Skidmore et al., 1990).

1.2.4.2. Entérocolite ulcéro-nécrosante

L'EUNC (entérocolite ulcéro-nécrosante) est une maladie gastro-intestinale présente chez 7 % des nouveau-nés prématurés, dont les plus immatures présentent un risque accru : 11 % des enfants atteints présentent un poids extrêmement faible ; 4 % des enfants avec un poids faible (Guillet et al., 2006). Cinquante pour cent des enfants atteints d'EUNC requièrent une intervention chirurgicale, dont la mortalité associée est plus importante par rapport à ceux traités médicalement (35 % contre 20 %) (Rich & Stephen, 2017). Cependant, chez ces derniers, la survie dépend encore du degré de prématurité et du poids de naissance (Fitzgibbons et al., 2009 ; Lin & Stoll, 2006). L'EUNC est une nécrose de l'intestin causée par une inflammation et invasion bactérienne de la paroi intestinale (Thompson & Bizzarro, 2008). Le début de l'EUNC est souvent insidieux, mais sa progression est rapide et varie inversement avec l'âge gestationnel (Gregory et al., 2011 ; Rich & Stephen, 2017). Bien que l'étiologie de l'EUNC reste incertaine, trois

facteurs de risque sont identifiés. Premièrement, une muqueuse perméable qui facilite la pénétration des bactéries pathogènes. Deuxièmement, l'intestin immature est passible face aux lésions ischémiques qui déclenchent une inflammation. Troisièmement, l'alimentation apportée par sonde gastrique altère l'intégrité de la muqueuse intestinale et conduit à une prolifération bactérienne (Eaton, Rees, & Hall, 2017; Gregory et al., 2011; Shulhan, Dicken, Hartling, & Larsen, 2017; Thompson & Bizzarro, 2008). Le diagnostic ainsi que la gravité de l'EUNC se base sur des signes cliniques (ex. un abdomen tendu) et des examens radiologiques (Neu, 2014). L'EUNC est également liée à la progression de lésions cérébrales (Divyen et al., 2008; Hall et al., 2008; Maitre, Marshall, Goldstein, Slaughter, & Price, 2011). L'EUNC comporte un risque neurodéveloppemental qui varie selon la gravité et l'âge gestationnel. Ainsi, les enfants traités chirurgicalement présentent un développement neurologique plus problématique que ceux traités médicalement (Allendorf et al., 2018 ; Hintz et al., 2005 ; Sonntag et al., 2010 ; Soraisham, Amin, Al-Hindi, Singhal, & Sauve, 2006). Cependant, indépendamment de sa gravité, l'EUNC constitue un risque accru de déficiences neurodéveloppementales chez les enfants avec un poids de naissance extrêmement faible (Shah et al., 2012). Ces déficiences impliquent des difficultés visuelles, cognitives et neuromotrices (Rees, Pierro, & Eaton, 2007 ; Schulzke, Deshpande, & Patole, 2007).

1.2.4.3. Sepsis

L'infection néonatale est une complication qui, tous les ans, cause environ 1.6 million de décès dans le monde. Le sepsis en est le premier responsable (Blencowe, Cousens, et al., 2013 ; Stoll et al., 2002 ; Stoll, 1997 ; Vergnano, Sharland, Kazembe, Mwansambo, & Heath, 2005). Le sepsis se développe chez 1.7 million de nouveau-nés prématurés, parmi lesquels l'incidence de cette maladie varie entre 21 % et 60 % selon le degré d'immaturité et la fragilité de l'enfant (Lawn et al., 2013). Le sepsis atteint 36 % des nouveau-nés extrêmement prématurés et environ 60 % de

ceux avec un poids extrêmement faible. Malgré les progrès pédiatriques, le taux du sepsis reste stationnaire et la mortalité associée peut aller jusqu'à 50 % (Lukacs & Schrag, 2012). Le sepsis se définit comme une inflammation générale et violente de l'organisme, en réponse à une infection invasive qui survient durant la période périnatale ou néonatale (Stoll et al., 2010). Cette maladie se diagnostique par l'observation des signes cliniques et un résultat positif de culture de sang pour la classer en fonction de son apparition (Stoll et al., 2002 ; Wynn, 2016). Le sepsis néonatal se classe en : apparition précoce et apparition tardive. Le sepsis d'apparition précoce est dû aux bactéries contractées avant ou pendant l'accouchement, pour lesquelles l'examen de culture est positif à moins de 72 heures de vie. Par conséquent, l'agent pathogène provient de la mère (Ganatra, Stoll, & Zaidi, 2010; Shane, Sánchez, & Stoll, 2017; Wynn et al., 2011). Le sepsis d'apparition tardive est dû aux bactéries contractées après l'accouchement, à plus de 72 heures de vie, dont l'agent pathogène provient de la communauté (Khan, Morris, & Bhutta, 2017). Le sepsis représente un risque neurologique chez les enfants nés prématurément, notamment parmi ceux avec un poids de naissance extrêmement faible, car il est associé avec des perturbations cognitives, sensorielles et motrices (Adams-Chapman et al., 2013; Bassler et al., 2010 ; Stoll, 2004 ; van der Ree, Tanis, Van Braeckel, Bos, & Roze, 2011 Ferreira, Mello, & Silva, 2014). Le sepsis est également corrélé avec des lésions cérébrales qui provoqueraient une diplégie, c'est à dire, un type de PC caractérisé par l'atteinte des deux membres inférieurs (Krallis et al., 2007; O'Shea, Klinepeter, Meis, & Dillard, 1998; Stoll, Hansen, Fanaroff, et al., 2004). Cela s'explique par le fait que la réponse inflammatoire secondaire à l'infection déclenche la production de cytokines (protéines qui régulent l'activité cellulaire) qui lèsent la substance blanche, et par la suite, provoquent des troubles neuromoteurs (Bashiri, Burstein, & Mazor, 2006; Dammann, Kuban, & Leviton, 2002; Murthy & Kennea, 2007).

1.2.4.4. Lésions cérébrales

Les lésions cérébrales sont une complication de la prématurité qui entraîne des anomalies de la substance blanche et grise. Appelées encéphalopathie de la prématurité, il s'agit d'anomalies de l'unité axon — soma qui sont associées aux troubles du développement. Les deux principales lésions cérébrales sont l'hémorragie intraventriculaire (HIV) et la leucomalacie périventriculaire (LPV) (Volpe, 1997 ; Ozen, & Burd, 2018). L'HIV et la LPV sont secondaires à des troubles de la circulation sanguine et processus infectieux (Volpe, 2009).

Hémorragie intraventriculaire

L'HIV (hémorragie intraventriculaire) est une complication qui atteint jusqu'au 45 % des nouveau-nés extrêmement prématurés ou avec un poids très faible, dont 15 % souffrent du degré le plus sévère. Ainsi, l'HIV survient principalement chez les enfants nés à 32 SA ou moins, et son incidence diminue avec l'augmentation de l'âge gestationnel (Ballabh, 2010). L'HIV est un saignement cérébral qui pénètre dans les ventricules, et qui commence dans la région périventriculaire nommée zone germinative. La zone germinative est un tissu transitoire à l'intérieur duquel se produit la prolifération des cellules souches neurales. La zone germinative est susceptible de subir un accident hémorragique, car elle possède un système vasculaire dense, composé de vaisseaux fragiles qui la rendent très sensible aux fluctuations du flux sanguin cérébral (McCrea et al., 2010). En d'autres termes, puisque le nouveau-né prématuré a des difficultés à maintenir un flux sanguin cérébral constant, les changements de la pression artérielle entraînent la rupture des dits vaisseaux. L'HIV se catégorise en quatre stades qui augmentent selon la gravité. Il semblerait que la gravité de l'HIV détermine le neurodéveloppement. Malgré des données conflictuelles, la plupart des études s'accordent à conclure que l'HIV I — Il s'associe à un risque accru de perturbations psychomotrices et cognitives

(Mukerji, Shah, & Shah, 2015 ; Volpe, 1997). En effet, les nouveau-nés prématurés avec HIV I — II présentent une probabilité importante d'anomalie neuromotrice grave par rapport à ceux nés à terme (soit 8 % contre 3 % respectivement) (Merhar, Tabangin, Meinen-Derr, & Schibler, 2012). Ces résultats s'avèrent encore plus importants chez les enfants les plus immatures (moins de 32 SA) dont 50 % présentent HIV ou LPV (Beaino et al., 2010; Patra, Wilson-Costello, Taylor, Mercuri-Minich, & Hack, 2006). De son côté, l'HIV III — IV s'associe à un risque de troubles neuromoteurs sévères, notamment la PC (Adams-Chapman, Hansen, Stoll, & Higgins, 2008 ; Pfahl et al., 2018). En effet, l'HIV la plus sévère est liée à des anomalies de la substance blanche périventriculaire, lesquelles sont fortement corrélées aux développements neurologiques les plus problématiques chez l'enfant né prématurément.

Leucomalacie périventriculaire

La LPV (leucomalacie périventriculaire) est la lésion cérébrale qui présage le plus une atteinte neurologique chez 15 % des nouveau-nés extrêmement prématurés (Volpe, 1998 ; Younge et al., 2017). La LPV consiste en une anomalie des axones qui forment la substance blanche périventriculaire en se caractérisant par deux composantes : une lésion diffuse et une lésion focale. La lésion diffuse consiste en la perte de processus ou la mort cellulaire et en la formation de gliose (cicatrices). En fonction de la composante focale, la LPV se catégorise en kystique et en non kystique. Dans la première, les nécroses focales présentent une taille macroscopique et évoluent en kystes (perte de tissu cérébral). Dans la seconde, les lésions nécrotiques sont microscopiques et évoluent en gliose (Volpe, 2009). Le mécanisme sous-jacent de la LPV est lié à deux facteurs : l'immaturation des artères cérébrales et la difficulté du système cérébrovasculaire à réguler les fluctuations du flux sanguin. Cela signifie qu'en réponse à une hypotension systémique, les nouveau-nés prématurés ne peuvent pas augmenter la perfusion cérébrale pour garantir le débit sanguin adéquat. Cette condition provoque un accident hypoxique

— ischémique qui produit une lésion cérébrale (Blumenthal, 2004). L'inflammation associée déclenche une cascade d'évènements qui lèsent le tissu et perturbent le développement des oligodendrocytes (cellules responsables de la production de la gaine de myéline qui entoure les axones). En d'autres termes, la LPV implique une déficience en oligodendrocytes, une hypomyélinisation axonale et une perte de volume cérébral (Khwaja et Volpe, 2008). La LPV se classe en quatre niveaux de gravité qui sont corrélés au développement (Novak et al., 2018). Dans ce sens, la plupart des enfants qui ont une LPV I présente un développement psychomoteur normal, alors que la majorité des enfants avec une LPV II — III présente un devenir développemental défavorable, tel que le diagnostic de PC (Beaino et al., 2010; Sarkar et al., 2018). Cela s'aggrave chez les enfants nés à moins de 32 SA qui présentent une LPV III — IV (Imamura et al., 2013). Étant donné que la LPV s'accompagne d'anomalies de la substance grise, sa gravité est également corrélée à des perturbations cognitives et sensorielles (Anderson et al., 2017 ; Hintz et al., 2005 ; Stoll, Hansen, Adams-Chapman, et al., 2004).

Bien que La LPV et l'HIV coexistent, elles sont deux entités différentes. Ainsi, l'HIV s'observe très précocement, alors que la LPV se détecte plus tardivement. La LPV est une lésion bilatérale et l'HIV est unilatérale. La LPV est corrélée à une diplégie, alors que l'HIV est liée à une hémiparésie. De même, elles diffèrent quant au diagnostic, car l'HIV est bien détectée par l'échographie cérébrale, alors que les composantes lésionnelles de la LPV s'avèrent mieux détectées par IRM.

1.3. Évaluation de la structure cérébrale

L'évaluation de la structure cérébrale, qui permet la détection précoce des lésions cérébrales, est fondamentale dans la prise en charge du nouveau-né prématuré. Cette évaluation se base sur des techniques d'imagerie dont les plus utilisées sont l'échographie cérébrale et l'IRM (imagerie par résonance magnétique). Ces techniques présentent des vertus et des inconvénients, en même temps qu'elles

différent quant au contexte d'utilisation, le maniement, et les résultats produits. D'une part, l'échographie cérébrale présente un coût relativement bas et elle peut s'utiliser au lit, même lorsque l'enfant est en état instable. De plus, l'échographie cérébrale ne comporte aucune radiation et permet une imagerie en série. Traditionnellement, les images s'obtiennent par la fontanelle antérieure. N'étant pas encore soudée, elle permet la visualisation des structures cérébrales importantes (ex. Les ventricules). D'autre part, l'IRM fournit des images *in vivo* et d'une haute qualité sur l'anatomie, le volume, la pathologie et la physiologie cérébrale. Cependant, l'utilisation de l'IRM chez le nouveau-né prématuré est limitée en raison des problèmes de logistique et de sécurité : il faut transporter l'enfant à la cabine d'évaluation ; des malaises mineurs après l'évaluation avec l'IRM sont fréquents (Heemskerk et al., 2012). L'IRM s'utilise donc principalement chez des enfants stables et proches du terme (Ecury-Goossen, Camfferman, Leijser, Govaert, & Dudink, 2015). L'utilisation de l'échographie cérébrale et de l'IRM dépend également de la nature des lésions cérébrales. Étant donné que l'HIV se développe les premiers jours de vie, l'« American Academy of Neurology » recommande que l'enfant prématuré bénéficie d'une échographie cérébrale entre le 7^e et le 14^e jour de vie (Allan & Volpe, 1986). En outre, comme il faut plus de temps pour identifier les changements de la substance blanche, des échographies répétées entre 36 SA à 40 SPT sont conseillées (de Vries, 1996). Cependant, comme la LPV la plus grave (kystique) est mieux détectée par l'IRM, le consensus en pédiatrie suggère un suivi par IRM entre 35 SA et 42 SPT (Khwaja & Volpe, 2008 ; McCrea et al., 2010). Néanmoins, en raison des contraintes de maniement, l'IRM ne s'utilise pas couramment dans un contexte clinique, bien que sa valeur prédictive des troubles neurodéveloppementaux soit avérée.

L'imagerie se destine au diagnostic des lésions cérébrales, elle est pourtant également utilisée pour le pronostic neurologique. De ce point de vue, les résultats obtenus par l'échographie cérébrale diffèrent de ceux obtenus par l'IRM. Ainsi, l'échographie cérébrale présente une sensibilité variable pour le pronostic des troubles cognitifs et neuromoteurs : 0.76 — 0.86 pour des enfants à plus de 32 SA ; 0.23 — 0.26 pour les enfants à moins de 32 SA. Concernant la prédiction de la PC,

la sensibilité de l'échographie cérébrale oscille entre 0.29 et 0.76. Cependant, l'échographie cérébrale présente une spécificité haute (0.86 — 0.95) pour prédire le développement typique. Concernant l'IRM, cette technique présente une sensibilité qui varie entre 0.71 et 0.94 pour prédire des troubles cognitifs et moteurs, alors qu'une haute sensibilité (0.82) pour prédire la PC. Malgré cela, la capacité de l'IRM à prédire le développement typique varie largement, car sa spécificité oscille entre 0.30 et 0.91 (de Vries, Van Haastert, Rademaker, Koopman, & Groenendaal, 2004; Mirmiran, 2004 ; Koivisto et al., 2010 ; Nongena, Ederies, Azzopardi, & Edwards, 2010 ; O'Shea et al., 2010; Herskind, Greisen, & Nielsen, 2014; Woodward, Anderson, Austin, Howard, & Inder, 2014).

Le fait que l'imagerie varie quant à l'estimation du risque neurologique met en avant l'importance de l'évaluation neurologique, car : elle permet d'évaluer la fonction cérébrale ; elle est plus sensible que l'échographie cérébrale et l'IRM pour pronostiquer un développement atypique ; combinée avec l'imagerie cérébrale, elle améliore la prédiction du développement (Bosanquet, Copeland, Ware, & Boyd, 2013 ; Constantinou, Adamson-Macedo, Mirmiran, & Fleisher, 2007 ; Maitre, Nelin, Noritz, Chorna, Williams, Carey, 2017). Ensemble, ces constats suggèrent qu'une approche holistique pour détecter les lésions cérébrales et les troubles neurologiques est préférable à un seul paradigme d'évaluation. C'est pourquoi l'évaluation neurologique, additionnée à l'imagerie cérébrale, représente un élément essentiel pour la prise en charge des nouveau-nés prématurés. De même, l'évaluation neurologique est une méthode peu invasive et peu coûteuse.

Encadré 1. Résumé du chapitre I

La naissance prématurée (avant 37 SA) est un problème de santé planétaire qui implique des séquelles neurologiques. Actuellement, la dichotomie entre prématuré et terme se remplace par la notion qui considère l'âge gestationnel comme un continuum dès le plus immature jusqu'au plus mature, dont le risque et la gravité des séquelles neurologiques augmentent avec la diminution de l'âge gestationnel. À cela s'agrègent les complications cliniques du nouveau-né prématuré qui exacerbent la vulnérabilité neurologique : la dysplasie bronchopulmonaire, l'entérocolite ulcéro-nécrosante, les infections et les lésions cérébrales. Dans ce sens, l'identification précoce des lésions cérébrales et l'évaluation du risque neurologique associé représentent un pilier dans l'intervention opportune du nouveau-né prématuré. L'intégrité cérébrale chez le nouveau-né prématuré se mesure par des techniques d'imagerie cérébrale, notamment l'échographie cérébrale dont l'utilisation est largement répandue et l'imagerie par résonance magnétique dont l'utilisation est limitée. Bien que l'imagerie se destine au diagnostic des lésions cérébrales, elle est également utilisée pour le pronostic neurologique. Cependant, il s'avère que la précision de ces techniques quant à l'estimation du risque neurologique varie largement. De ce fait, bien qu'il existe des variables biologiques (ex. L'âge gestationnel), des complications (ex. Lésion cérébrale) et des biomarqueurs (ex. Signes moteurs), ces facteurs, à eux seuls, ne permettent pas d'identifier les enfants à risque neurologique. C'est ainsi que le besoin d'outils d'évaluation neurologique appropriés pour le nouveau-né prématuré reste une question pertinente. Ces aspects seront abordés dans le chapitre suivant.

II. L'évaluation neuromotrice

2.1. Évaluation de la fonction cérébrale

L'évaluation de la fonction cérébrale, qui permet la détection précoce des signes de dysfonction, est fondamentale dans la prise en charge du nouveau-né prématuré. Cependant, la fonction neurologique est un phénomène non directement observable, donc non directement mesurable. La psychométrie introduisit le concept « construit »² pour désigner un phénomène non observable, mais qui est défini et délimité dans un cadre scientifique (de Vet, Terwee, Mokkink & Knol, 2011). Bien qu'un construit ne soit pas directement observable, il est susceptible d'être évalué par des caractéristiques observables qui lui sont associées. Par conséquent, la fonction neurologique est un construit non directement mesurable, mais qui peut être évalué par le comportement moteur observable (Majnemer & Mazer, 1998). Ainsi, la neuromotricité constitue le répertoire moteur, observable et mesurable, qui reflète le fonctionnement neurologique. C'est pourquoi l'« American Academy of Pediatrics » recommande que l'enfant né prématurément bénéficie d'une évaluation neuromotrice adaptée, précoce, opportune et longitudinale, afin de déterminer une intervention précoce (Wang, Mcglynn, Brook, Leonard, & Piecuch, 2006; Lipkin, 2005). Cela est crucial, car l'enfant prématuré profite davantage d'une intervention lorsqu'elle démarre à la période néonatale, et ce quand la plasticité cérébrale est la plus prolifique (Blauw-Hospers & Hadders-Algra, 2005; Ismail, Fatemi, & Johnston, 2017). L'évaluation neuromotrice est fondamentale dans les

² De l'anglais « Construct »

soins au nouveau-né prématuré, car elle permet d'accomplir trois objectifs cliniques : distinguer l'enfant ayant un développement neuromoteur typique et celui ayant des signes pathologiques ; prédire le développement dans le futur sur la base de la motricité actuelle ; évaluer le changement neuromoteur au cours du temps (Kirshner & Guyatt, 1985). Plus particulièrement, les mesures neuromotrices comportent une valeur clinique pour : réaliser un diagnostic différentiel ; localiser une lésion du système nerveux ; caractériser la dysfonction neurologique ; suivre l'évolution d'une lésion ; estimer l'impact des interventions ; orienter les parents (Lipkin, 2005 ; Majnemer & Mazer, 1998 ; Smoliar, 1989 ; Wusthoff, 2013). Néanmoins, l'évaluation neuromotrice doit relever plusieurs défis provoqués par la prématurité. Premièrement, les références normatives doivent être spécifiques de l'enfant prématuré, car sa trajectoire développementale diffère de celle de l'enfant né à terme. Deuxièmement, les tests neuromoteurs doivent refléter les différences maturationnelles, car, par exemple, un cerveau à moins de 34 SA diffère d'un cerveau à plus de 34 SA. Troisièmement, l'évaluation se doit de représenter la trajectoire évolutive de la motricité, car des changements cérébraux surviennent à différents moments. Quatrièmement, les mesures doivent détecter les troubles et les retards le plus précocement possible, car la période entre 28 SA et 15 mois de vie offrirait les meilleures opportunités d'intervention. Cinquièmement, les mesures longitudinales sont préférables à celles transversales, car elles fournissent plus d'information. Sixièmement, l'évaluation doit éviter la manipulation excessive de l'enfant, car elle refléterait le stress, mais pas l'intégrité neurologique (Sweeney & Blackburn, 2013; Back, Riddle, & McClure, 2007; Hadders-Algra, 2014; Gosselin, Gahagan, & Amiel-Tison, 2005).

2.2. Les outils d'évaluation neuromotrice

Les outils d'évaluation neuromotrice se classent en quatre catégories : l'examen neurologique qui évalue les réponses sensorimotrices, la posture et le tonus ; les

méthodes qui allient l'examen neurologique et l'observation du comportement ; l'observation des habiletés motrices selon le stade développemental ; l'observation de la qualité des patterns moteurs (Tableau 1). Bien évidemment, la nature d'un test neuromoteur s'explique par l'approche à partir de laquelle il fut conçu (Lipkin, 2005). Ainsi, deux approches du développement moteur sont identifiées. L'approche classique considère que le développement moteur se produit d'une manière linéaire et invariable, grâce au contrôle inhibitoire qui croît avec la maturation corticale (Ballabh, 2010 ; Case-Smith, 1996 ; Dalton, 2005). L'approche contemporaine considère que la variabilité (génétique et épigénétique) est cruciale pour le développement, car elle permet la sélection des patterns moteurs les plus adaptés, et ce par l'interaction entre des composants endogènes et environnementaux (Turvey, 1990 ; Thelen, 1995 ; Sporns & Edelman, 1993 ; Edelman, 1993 ; Hadders-Algra, 2010, 2018a). C'est la raison pour laquelle les tests classiques (ex. le Bayley) mesurent l'acquisition d'habiletés motrices selon un ordre chronologique précis, alors que les tests contemporains (ex. l'Infant Motor Profile) évaluent la qualité des facteurs qui déterminent une compétence motrice (ex. la variabilité) (Case-Smith, 1996 ; Long & Tieman, 1998). En somme, bien que l'évaluation classique et celle contemporaine mesurent la motricité observable, elles diffèrent quant aux composants considérés. Cela détermine le cadre d'utilisation, l'utilité clinique, l'objectif d'évaluation et l'adéquation à une population particulière. Dans cette mesure, deux tests s'avèrent être les plus adaptés pour le nouveau-né prématuré : Le Test of infant motor performance (TIMP) et l'analyse des GMs (mouvements généraux).

Tableau 1. Mesures neuromotrices pour l'enfant né prématurément

Instrument	Âge	Objectif	Composants évalués	Variable	Évaluateur (Durée en minutes)
<i>Examen neurologique</i>					
Touwen	0 mois — 18 mois	D	Nerfs crâniens, réflexes, réactions, coordination, motricité fine et globale	Cat	Professionnels en développement neuromoteur (15)
Amiel-Tison	0 mois — 6 ans	D	Tonus musculaire, nerfs crâniens, développement moteur, motricité spontanée, réflexes, réactions	Cat	Professionnels en développement neuromoteur (10)
Active and Passive muscle power	3 mois — 12 mois	D, P	Puissance musculaire active et passive, stabilité posturale, fluidité motrice	Cat	AD
Hammersmith infant neurological examination	2 mois — 24 mois	D, P	Nerfs crâniens, posture, tonus, réflexes, réactions, développement moteur, états comportementaux	Cont	Staff médical sans expérience particulière (AD)
<i>Procédures avec codage normalisé</i>					
Primitive reflex profile	0 mois — 2 ans	D	9 Reflexes archaïques	Cont	Pédiatres (AD)
Infant neurological international battery	1 mois — 18 mois	D	Spasticité, fonction vestibulaire, contrôle de la tête/tronc, tonus, comportement moteur	Cat	Staff médical sans expérience particulière (AD)

des jambes

Bayley scales of infant development	1 mois — 3 ans	D, E	Échelle motrice : motricité globale et fine ; échelle mentale ; échelle comportementale	Cont	Psychologues (25 – 60)
Peabody developmental motor scales	0 mois — 6 ans	D	Motricité globale et fine ; réflexes, locomotion, manipulation d'objets, intégration visuomotrice	Cont	Ergothérapeutes (45 – 60)
Movement assessment of infants	0 mois — 12 mois	D, E	Tonus musculaire, réflexes, réactions automatiques et comportement volitif.	Cont	Professionnels en développement neuromoteur (20 – 30)
Neuromotor behavioral inventory	0 mois — 12 mois	D	Tonus musculaire, développement moteur, qualité de la motricité, réflexes, réactions, comportement oral	Cont	AD
Toddler and infant motor evaluation	4 mois – 3,5 ans	D, E	Organisation motrice ; habiletés sociales/émotionnelles ; performance fonctionnelle ; aspects cliniques	Cont	Professionnels experts en l'évaluation motrice (10 – 55)
<i>Observation du comportement moteur</i>					
Alberta infant motor scale	0 mois — 18 mois	D	Position posturale : couché sur le dos et sur le ventre ; assis et debout	Cont	Pédiatres et kinésithérapeutes (15)
Structured observation of motor performance	0 mois — 10 mois	D	Qualité de la posture et de la motricité	Cont	Professionnels en développement neuromoteur (AD)

Qualité du comportement moteur

Test of infant motor performance	32 SA – 4 mois	D	Contrôle moteur et organisation posturale ; activités fonctionnelles ; réactions ; mouvements spontanés	Cont	Kinésithérapeutes et ergothérapeutes (AD)
Mouvements généraux	Préterme — 5 mois	D, P	Variabilité et complexité du comportement moteur spontané couché sur le dos	Cat	Aucune profession spécifique, mais formation exigée (3)

m, mois ; D, discriminatif ; Cat, donnée catégorielle ; P, prédictif ; AD, aucune donnée ; Cont, donnée continue ; E, évaluatif. *Issu, traduit et adapté de Heineman & Hadders-Algra, (2008)*

2.2.1. Les outils d'évaluation neuromotrice destinés au nouveau-né prématuré

Le Test of infant motor performance (TIMP) et l'analyse des GMs (mouvements généraux) sont les instruments les mieux adaptés au nouveau-né prématuré, car ils permettent d'identifier les enfants qui pourraient bénéficier d'une intervention (Heineman & Hadders-Algra, 2008 ; Spittle et al., 2008 ; Noble & Boyd, 2012 ; Craciunoiu & Holsti, 2016 ; Peyton & Einspieler, 2018 ; Song et al., 2018). Bien que le TIMP nécessite une manipulation de l'enfant, alors que l'analyse des GMs n'en implique pas, il est intéressant de préciser que ces deux tests évaluent le même composant : la motricité spontanée. Nous allons les décrire plus en détail.

2.2.1.1. Le Test of Infant Motor Performance

Le TIMP (Test of Infant Motor Performance) évalue la qualité du comportement moteur, afin de détecter les enfants à risque développemental entre 32 SA et 4e mois de vie. Le TIMP ne requiert pas de cours de formation, mais il doit être appliqué par des kinésithérapeutes ou ergothérapeutes. Cet instrument comprend deux échelles dont le codage dure 36 minutes. L'échelle d'observation contient 13 items dichotomiques (typique/atypique) qui évaluent la qualité de la motricité spontanée. L'échelle des réponses sur sollicitation comprend 29 items qui évaluent, sur quatre à sept points, le contrôle moteur et l'organisation posturale quand l'enfant est placé dans plusieurs positions (ex. le contrôle de la tête en position assise) (Campbell & Kolobe, 2000). De même, le TIMP évalue le système neurologique et le système comportemental (ex. l'orientation vers les stimuli auditifs) (Campbell, Kolobe, Osten, Lenke, & Girolami, 1995). Les résultats du TIMP comprennent : le score brut qui indique les items correctement cotés ; le score corrigé en fonction de l'âge ; le

classement en centiles ; le score de développement. Ce test est normalisé sur un échantillon de 990 enfants nés aux États-Unis ayant des risques développementaux.

Le TIMPSI est une version courte du TIMP à des fins de dépistage. Il se compose de 29 des 42 items du TIMP, lesquels incluent l'observation des mouvements spontanés et les réponses produites par des stimuli sensoriels (Campbell, Swanlund, Smith, Liao, & Zawacki, 2008). Le TIMPSI présente une précision diagnostique haute (0.81). De son côté, pour pronostiquer un développement atypique à 6 mois et 5 ans, la sensibilité du TIMP fluctue considérablement entre basse (0.50) et haute (1.00), alors que sa spécificité pour prédire le développement typique se situe entre moyenne (0.66) et haute (1.00) (Craciunoiu & Holsti, 2016). De même, le TIMP démontre une bonne capacité à détecter des changements développementaux, car les scores augmentent significativement entre 32 SA et 36 SA (Craciunoiu & Holsti, 2016). Cependant, nous avons rencontré des difficultés pour trouver des études sur la capacité diagnostique du TIMP sur les nouveau-nés prématurés. En effet, le seul article trouvé aborde cette question avec plusieurs groupes (Rose & Westcott, 2005). De plus, une revue déclare que le TIMP manque de preuves quant à sa qualité psychométrique pour diagnostiquer des populations à risque, dont les enfants prématurés (Novak et al., 2017).

2.2.1.2. L'analyse des mouvements généraux

L'analyse des GMs (mouvements généraux) s'utilise pour identifier les enfants qui présentent un dysfonctionnement cérébral entre la période préterme et le 5e mois de vie. Les GMs ont été décrits comme un pattern qui émerge dans la vie foetale et intègre le répertoire moteur spontané (Prechtl, 2001). Les GMs consistent en des séquences motrices, globales et variables, où toutes les parties du corps y participent sans suivre un ordre distinctif (Einspieler et al, 2004). Bien que l'analyse des GMs ne requière pas de profession spécifique, une formation est nécessaire pour devenir un observateur qualifié. Ce test n'implique aucune manipulation,

puisque le codage s'effectue sur observation de quelques minutes d'enregistrement de vidéo des GMs quand l'enfant est couché sur le dos. Bien que les GMs de la période préterme puissent s'évaluer alors que l'enfant dort, l'analyse des GMs à partir du terme doit se réaliser dans un état d'éveil approprié. L'analyse des GMs considère trois aspects : la variabilité concernant la vitesse et la taille du répertoire moteur ; la complexité par rapport aux membres participants ; la fluidité déterminée par la souplesse. Une autre manière de coder les GMs ajoute plus de variables à considérer (Hadders-Algra et al., 1997). Indépendamment de l'approche d'évaluation, les GMs typiques exhibent une grande variabilité, alors que les mouvements atypiques présentent une variabilité réduite. Cependant, les GMs typiques et les GMs atypiques présentent des traits spécifiques du stade développemental (ces aspects seront approfondis dans le chapitre suivant). Une échelle permettant de quantifier l'analyse des GMs est également disponible (Einspieler et al, 2004). Cette échelle note entre 0 et 2 la qualité des GMs, et ce d'une manière détaillée pour les membres supérieurs et inférieurs. La présence de signes neurologiques (ex. jerks) et la qualité du répertoire moteur, entre le 3^e mois et le 5^e de vie, sont considérées (Einspieler et al., 2015). Enfin, cette analyse détaillée produit un score optimal des GMs qui permet de détecter des changements subtils et de prédire une détérioration à court terme.

Même si la majorité des études souligne la bonne (0.81 – 0.98) fiabilité intercodeur, d'autres études n'ont pas reproduit le même bon indicateur (Bernhardt, Marbacher, Hilfiker, & Radlinger, 2011; Fjørtoft, Einspieler, Adde, & Strand, 2009; Valentin, Uhl, & Einspieler, 2005). En effet, il semblerait que la fiabilité des GMs augmente avec l'âge de l'enfant au moment de l'observation. En ce qui concerne sa capacité diagnostique, les GMs présentent une sensibilité entre moyenne (0.62) et haute (1.00) pour prédire un devenir développemental atypique à 18 mois de vie (Craciunoiu & Holsti, 2016). Cependant, la spécificité pour prédire le développement typique est entre basse (0.44) et moyenne (0.70). En fait, la valeur prédictive des GMs varie en fonction de l'âge corrigé et le type de devenir. Spécifiquement, la meilleure prédiction s'obtient par une série longitudinale d'évaluations, alors que la prédiction d'une évaluation unique s'améliore avec l'âge, à partir de 35 SA. La

persistance d'anomalies motrices entre 35 SA et au-delà de 38 SA augmente l'aptitude des GMs (0.80 — 1.00) à prédire un développement atypique (Craciunoiu & Holsti, 2016). L'évaluation des GMs à la période post terme, entre le 3e mois et le 5e mois de vie, est particulièrement sensible (0.95 — 1.00) à prédire la PC (Darsaklis, 2011). Cependant, la validation systématique de l'analyse qualitative et l'analyse quantitative des GMs, réalisée chez le nouveau-né prématuré, demeurent encore un sujet à explorer.

Malgré la pertinence clinique de l'analyse des GMs, son utilisation n'est pas répandue dans les services de néonatalogie. Les difficultés suivantes peuvent se présenter : la contrainte de faire coïncider des conditions optimales chez l'enfant avec la disponibilité d'un observateur qualifié ; le coût de déplacement du personnel pour aller se former auprès de l'équipe qui a développé cette technique ; le risque d'une interférence subjective liée à l'observateur pouvant fausser les résultats de l'évaluation. (Adde, Helbostad, Jensenius, Taraldsen, & Støen, 2009; Marcroft, Khan, Embleton, Trenell, & Plötz, 2015). En dépit de ces difficultés et en vertu de son utilité clinique, l'analyse des GMs, aussi bien que le TIMP, reste une des techniques les mieux adaptées aux besoins des nouveau-nés prématurés. Étant donné que ces deux mesures évaluent la motricité spontanée, le chapitre suivant approfondit les bases neurologiques du développement et fonctionnement de la motricité spontanée, ainsi que la pertinence clinique des GMs.

Encadré 2. Résumé du chapitre II

En étant un attribut non directement observable, la fonction neurologique s'évalue en analysant le répertoire moteur. Ainsi, la fonction neuromotrice se définit comme l'ensemble de mouvements observables, donc mesurables, qui rendent compte de l'intégrité neurologique chez le nouveau-né prématuré. Dans cette mesure, les outils qui évaluent la neuromotricité du nouveau-né prématuré s'avèrent à la fois primordiaux et complexes : primordiaux, car ils fournissent les bases pour l'intervention précoce ; complexes, car ils doivent s'adapter aux particularités d'un enfant fragile. Cela peut expliquer le fait qu'il n'existe que deux tests neurologiques adaptés au nouveau-né prématuré : le TIMP (Test of Infant Motor Performance) et l'analyse des GMs. D'une part, le TIMP comporte un degré de manipulation de l'enfant et manque d'une validation systématique sur le nouveau-né prématuré. D'autre part, l'analyse des GMs est la méthode qui implique la moindre manipulation de l'enfant. Cependant, la validation systématique de l'analyse des GMs sur le nouveau-né prématuré reste encore à explorer et des contraintes liées à la formation du personnel, des facteurs logistiques et subjectifs pourraient limiter son utilisation. Cela implique le besoin de mesures neuromotrices validées, faciles à mettre en place et objectives, destinées au nouveau-né prématuré. En vertu de ses avantages et en dépit de ses inconvénients, le TIMP et l'analyse des GMs se rejoignent sur le fait d'évaluer la motricité spontanée. C'est pourquoi le chapitre suivant présente les bases neurologiques de la motricité spontanée, en se focalisant sur les GMs et leur pertinence clinique pour l'enfant né prématurément.

III. Les mouvements généraux

3.1. Parcours évolutif

La motricité fœtale, émergeant sans l'intervention d'aucun stimulus apparent, permet le développement du système neuromoteur (Einspieler, Prayer, & Prechtl, 2012). De même, elle rend compte de l'intégrité d'un jeune système nerveux (Einspieler & Prechtl, 2005 ; Kurjak et al., 2008 ; Kurjak, Stanojevic, Salihagic-kadic, & Azumendi, 2004). Ainsi, les premiers mouvements consistent en des flexions latérales et simples de la tête ou tronc qui commencent à 7 SA en coïncidant avec la formation des premières synapses dans la moelle épinière (Okado, 1980). Ces mouvements, à peine perceptibles, impliquent un déplacement lent et stéréotypé des contours du fœtus. Le deuxième pattern moteur à apparaître est le sursaut. Visible à 7.5 SA — 8 SA, le sursaut est un mouvement généralisé, rapide et simple. Ensuite, les GMs émergent comme le premier pattern variable à 9 SA — 10 SA. À cet âge, les voies spinales ne sont pas encore terminées, les premières synapses corticales apparaissent, les mouvements à peine perceptibles commencent à disparaître et la fréquence des GMs augmente (de Vries, Visser, & Prechtl, 1982; Lüchinger, Hadders-Algra, van Kan, & de Vries, 2008 ; Okado, 1980). Ces aspects de maturation reflètent trois caractéristiques du système moteur précoce : un contrôle sous-cortical primitif ; un contrôle cortical émergeant ; la prépondérance de la motricité spontanée dans le développement. Au deuxième trimestre, le répertoire moteur du fœtus se compose d'étirements, sursauts, rotations, flexions/extensions, ouverture de la mâchoire, succion, déglutition et GMs (Tableau 2). Au troisième

trimestre, certains mouvements sont encore stéréotypés et simples, alors que d'autres deviennent plus complexes (Okado, 1980). À la naissance, le répertoire moteur se caractérise par la coexistence de mouvements simples et complexes qui évoluent au cours du développement. Parmi ceux complexes, les GMs sont l'un des plus stables, car ils gardent leurs qualités constitutives (complexité, variabilité et fluidité) tout au long de leur parcours évolutif. Toutefois, les GMs présentent des caractéristiques spécifiques qui ont été décrites en trois étapes : préterme, writhing et fidgety.

Tableau 2. Premiers mouvements au cours du premier trimestre de gestation

Mouvements	Semaine d'aménorrhée
Activité cardiaque fœtale	5.5 - 6.5
Mouvements à peine discernables	7.5 - 8.5
Sursaut	8.0 - 9.5
GMs	8.5 - 9.5
Étirement	10.5 - 15.5
Rotation	10.0 - 11.0
Mouvements isolés bras/jambe	9.5 - 10.5
Ouverture de la mâchoire	10.5 - 12.5
Succion et déglutition	12.5 - 14.5
Bâillement	11.5 - 15.5
Mouvements respiratoires	10.5 - 11.5
Hoquet	8.5 - 10.5
Mouvements oculaires	
Lent	16.0
Rapide	23.0

Adapté et traduit de Kurjak et al., (2008)

Les GMs de la période préterme apparaissent à 10 SA, et évoluent en writhing à l'âge du terme. Dans la période préterme, les GMs se caractérisent par des flexions/extensions des membres avec une grande amplitude, une vitesse rapide et une participation notable du tronc et du pelvis. Les GMs de la période writhing émergent à terme, à 38 SA, et s'observent jusque 5 SPT (1er mois de vie). À terme,

les GMs présentent une séquence de contorsions elliptiques des bras, jambes et cou. Les GMs à terme sont plus puissants que ceux à préterme, mais avec une amplitude moins large, une vitesse plus lente et une moindre participation du tronc. Les GMs de la période fidgety émergent à 6 SPT (2e mois) et s'observent jusqu'à 22 SPT (6e mois). À la période post terme, les GMs constituent un flot de petits mouvements rythmiques qui surviennent sur le corps tout entier dont toutes les parties participent à importance égale. Nettement différents des writhing, les fidgety sont doucement arrondis, de faible amplitude, de vitesse modérée avec une accélération variable dans toutes les directions (Prechtl & Hopkins, 1986). Bien que les GMs présentent des différences évolutives, la variabilité (spatiale et temporelle) est la caractéristique transversale à tous les stades du développement (Figure 1). Cependant, dans le cadre d'une lésion ou dysfonction cérébrale, les GMs perdent leur variabilité distinctive et deviennent atypiques (Figure 2). Trois types de GMs atypiques s'observent à périodes préterme et terme : le « pauvre répertoire » caractérisé par des séquences répétitives ; le « chaotique » défini par des mouvements désorganisés ; le « cramped-synchronised » caractérisé par une grande rigidité. À post-terme, les « fidgety atypiques » sont soit exagérés soit intermittents.

3.2. Corrélats neuronaux

L'activité neuronale intrinsèque est le mécanisme fonctionnel des GMs, car elle est la base du fonctionnement cérébral et s'observe au niveau du cortex fœtal (Gusnard & Raichle, 2001; Hadders-Algra, 2018; Moore, Zhou, Jakovcevski, Zecevic, & Antic, 2011). Ainsi, les GMs commencent par des rafales de potentiels d'action générés spontanément au stade où des neurones constituent les circuits sous-corticaux qui produisent les GMs et le circuit corticospinal qui les module.

Figure 1. Variabilité des GMs



Variabilité des GMs chez un enfant âgé de 36 AGc. La figure se compose des images d'une minute d'enregistrement de vidéo avec un intervalle de 5 secondes entre chaque cadre.

Figure 2. Variabilité réduite des GMs



Variabilité réduite des GMs chez un enfant âgé de 35 AGc. La figure se compose des images d'une minute d'enregistrement de vidéo avec un intervalle de 5 secondes entre chaque cadre.

3.2.1. Générateurs centraux de patterns

À la période embryonnaire, les interneurones sous-corticaux forment les générateurs centraux de patterns (CPGs)³ qui produisent les GMs (Feller, 1999). Les CPGs sont responsables de : la production des mouvements rythmiques ; la coordination flexion/extension ; l'alternance droite/gauche (Baev & Shimansky, 1992; Zhong, Shevtsova, Rybak, & Harris-Warrick, 2012). Les preuves de l'existence des CPGs proviennent des modèles animaux. Ainsi, des chats avec dissection spinale et/ou corticale, placés sur un tapis roulant, présentent un comportement locomoteur (Barbeau & Rossignol, 1987 ; Forssberg, Grillner, Halbertsma, & Rossignol, 1980 ; Grillner & Zangger, 1984). De même, des motoneurones *in vitro* présentent une activité dans l'absence de tout mouvement effectif (Grillner & Zangger, 1984) ; (Frigon, 2012). Bien que l'activité des CPGs soit intrinsèque, elle est modulée corticalement et périphériquement pour rendre le mouvement souple (Cohen, Guan, Harris, Jung, & Kiemel, 1996 ; Dietz, 2003 ; Mackay-Lyons, 2002). En effet, des chats décortiqués présentent des mouvements, mais ils manquent de souplesse (Grillner & Wallén, 1985 ; Martinez, Tuznik, Delivet-Mongrain, & Rossignol, 2013). Vu les caractéristiques des CPGs chez l'animal, ils sont proposés comme le substrat neuronal des GMs (Mackay-Lyons, 2002 ; Grillner & Wallén, 1985 ; Martinez, Tuznik, Delivet-Mongrain, & Rossignol, 2013). Néanmoins, existe-t-il des CPGs chez l'humain ?

3.2.2. Générateurs centraux de patterns chez l'humain

Les preuves sur l'existence des CPGs chez l'humain sont indirectes, car elles se basent sur l'observation de patients. De ce fait, des patients avec des lésions de la moelle épinière présentent : l'alternance flexion/extension, des contractions

³ De l'anglais « Central Pattern Generators »

involontaires et l'alternance musculaire (agoniste/antagoniste). De même, des pas coordonnés ont été observés chez des patients paraplégiques placés sur un tapis roulant, ou soumis à des stimulations électriques dans la région lombo-sacrée (Bussel et al., 1988 ; Calancie, 2006 ; Nadeau, Jacquemin, Fournier, Lamarre, & Rossignol, 2010). L'électromyographie suggère que cette activité motrice était due à la libération partielle des CPGs situés dans la colonne vertébrale, notamment dans la région lombo-sacrée. Ces données démontrent l'existence des GPCs capables de générer des mouvements spontanés chez l'humain (Danner et al., 2015 ; Minassian, Hofstoetter, Dzeladini, Guertin, & Ijspeert, 2017). Malgré la ressemblance des CPGs humain et animal, l'activité des CPGs humains requiert davantage de la rétroaction sensorielle et la modulation supraspinale (Calancie, 2006).

3.2.3. *Modulation*

L'activité des CPGs, se traduisant par les GMs, est modulée par l'activité de la sous-plaque qui se trouve dans la substance blanche (Hadders-Algra, 2007). La sous-plaque joue un rôle décisif dans le développement cortical, car elle : sert de compartiment d'attente pour la migration et perfectionnement des fibres corticales ; constitue le site principal de la synaptogenèse et différenciation neuronale ; règle la connectivité corticale (Volpe, 2009). Par rapport au développement de la sous-plaque : elle commence son activité synaptique à 9 — 10 SA ; elle augmente son volume et la synaptogenèse entre 28 et 34 SA ; elle diminue progressivement après 34 SA ; elle reste présente jusqu'à la période postnatale (Kostovic, Goran, Vulksic, & Judas, 2015). Plusieurs arguments étayent l'hypothèse selon laquelle la sous-plaque module les GMs, car elle confère aux GMs leur complexité et leur variabilité : les premières synapses dans la sous-plaque coïncident avec l'émergence des GMs à 9 — 10 SA ; la présence de la sous-plaque correspond à la présence des GMs entre 9 — 10 SA et 22 SPT (\approx 6e mois) ; les fœtus anencéphaliques exhibent des

GMs, mais atypiques ; les embryons de poulet n'ont pas de sous-plaque et de mouvements complexes, alors que les embryons des mammifères détiennent une sous-plaque et des mouvements complexes. À terme (40 SA), l'apparition des GMs writhing coïncide avec : la diminution du circuit moteur transitoire, concentré dans la sous-plaque ; l'expansion du circuit moteur permanent, aussi nommé corticospinal. Cependant, la modulation des GMs s'effectue davantage par le circuit corticospinal (Hayat & Rutherford, 2018). À post terme (6 SPT et 22 SPT), l'apparition des GMs fidgety coïncide avec la disparition du circuit transitoire et l'augmentation des inputs sensoriels vers le circuit corticospinal. De ce fait, les GMs fidgety révèlent une activité motrice de moins en moins modulée par le cortex et modulée davantage par l'information sensorielle (Rubinstein et al., 1989 ; Hadders-Algra, 2007). Cette modulation, corticale et sensorielle, induit la variabilité et la complexité des GMs. C'est la raison pour laquelle, lors d'une anomalie cérébrale, les GMs deviennent atypiques, c'est à dire, leur complexité et leur variabilité diminuent.

3.2.4. Importance clinique

Les GMs atypiques reflètent les anomalies cérébrales qui surviennent chez l'enfant né prématurément (Einspieler, 2008). À préterme, les GMs atypiques sont un marqueur de difficultés maturationnelles telles que la diminution du volume de la substance grise et du cervelet (George et al., 2018 ; Olsen et al., 2016). En effet, la perte de volume cérébral et cérébelleux est associée aux troubles développementaux sévères à 1 an, ainsi qu'au retard moteur et cognitif à 2 ans (Inder, 2005 ; Park et al., 2014). À terme, les GMs atypiques reflètent des lésions et des perturbations maturationnelles telles que : diminution du volume de la substance blanche ; dilatation ventriculaire ; amincissement du corps calleux ; LPV (leucomalacie périventriculaire) (Ferrari et al., 2002). Les GMs atypiques à terme sont associés aux troubles comportementaux et cognitifs. En effet, les enfants ayant des GMs atypiques qui persistent au-delà de 2 mois de vie présentent un quotient

intellectuel plus faible que les enfants avec une normalisation plus précoce des GMs (Kanemaru et al., 2013). Les GMs atypiques dénommés *cramped synchronized*, se caractérisant par une rigidité sévère, prédisent très bien la PC (Ferrari et al., 2002). À post terme, les GMs atypiques témoignent d'anomalies maturationnelles et lésionnelles telles que : perte de volume périventriculaire ; dilatation ventriculaire ; diminution du corps calleux ; LPV (Grunewaldt et al., 2014; Peyton et al., 2016; Spittle et al., 2008; van Iersel, Bakker, Jonker, & Hadders-Algra, 2010). Les GMs *fidgety atypiques* sont un indicateur prédictif de la PC, des troubles de la motricité fine et la motricité globale. De même, les *fidgety atypiques* sont associés aux troubles de la mémoire, l'intelligence et le langage (Bock, Will, Behrenbeck, & Jarczok, 2017; Cioni et al., 2000 ; Einspieler, 2008 ; Hamer, Bos, & Hadders-Algra, 2016 ; Skiöld, Eriksson, Eliasson, Ådén, & Vollmer, 2013 ; Grunewaldt et al., 2014 ; Peyton et al., 2017; Zang et al., 2016 ; Bruggink, Koenraad, Braeckel, & Bos, 2010 ; Butcher et al., 2009 ; Fjørtoft et al., 2013 ; Grunewaldt et al., 2014). La LPV est l'anomalie cérébrale la plus associée aux GMs atypiques chez l'enfant né prématurément. Cette pathologie, qui atteint la sous-plaque corticale et les neurones migratoires, perturbe la maturation des oligodendrocytes responsables de la myélinisation axonale (Back et al., 2007; Volpe, 2009). De ce fait, les troubles développementaux de l'enfant né prématurément reflèteraient un problème dans la transmission d'informations. Cela pourrait expliquer le fait que les GMs sont corrélés avec des anomalies dans les voies de projection tels la corona radiata, la capsule interne et le fascicule longitudinal (Galinsky et al., 2018).

Malgré son importance clinique, trois problèmes ne favorisent pas l'utilisation de l'analyse des GMs : la forte implication de la subjectivité du codeur ; le risque d'une fiabilité faible, notamment pour les observateurs moins formés ; le manque de précision concernant les paramètres temporeux (Marcroft et al., 2015) ; (Karch et al., 2008). Dans ce cadre, l'évaluation objective de la neuromotricité chez le nouveau-né prématuré reste une question pertinente, dont la réponse impliquerait l'utilisation de technologies de reconnaissance automatisée des GMs.

3.3. Évaluation automatisée

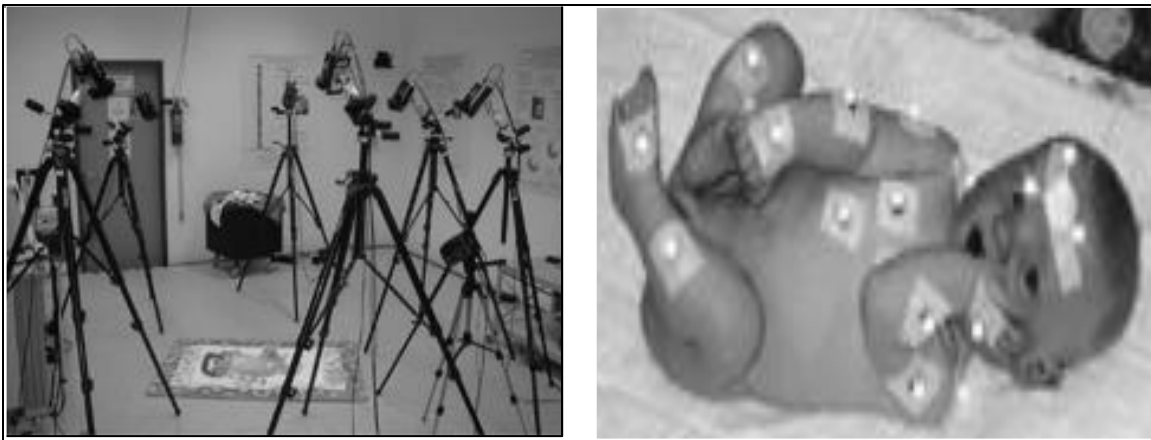
L'évaluation neuromotrice du nouveau-né prématuré représente un défi qui doit être relevé par l'adoption d'une approche pluridisciplinaire (Marschik et al., 2017). L'informatique, associée à la clinique, développe des systèmes automatisés qui, basés sur des modèles d'activité, analysent objectivement les GMs. Il s'agit de systèmes de reconnaissance automatisée qui se classent selon la méthode de détection utilisée : ceux indirects qui utilisent des dispositifs de vidéo ; ceux directs qui emploient des capteurs collés au corps (Marcroft et al., 2015).

3.3.1. Reconnaissance par détection indirecte

Les systèmes par détection indirecte utilisent des dispositifs de vidéo qui enregistrent les GMs soit en trois dimensions (3D), soit en deux dimensions (2D). Les systèmes basés sur une détection en 3D requièrent des caméscopes haut de gamme, ainsi que des marqueurs collés aux membres qui seront traqués (Figure 3). Ces systèmes ont permis l'identification des indicateurs d'anomalie suivants : l'indice jerk élevé (quantité de secousses) ; l'asymétrie de la vitesse des pieds ; la corrélation entre la trajectoire des membres (synchronisation motrice) ; l'aire de la trajectoire des pieds différant de la moyenne globale (mouvement non fluide) ; l'aire de la trajectoire des pieds hors de l'écart type de la moyenne globale (mouvement non harmonieux). (Kanemaru et al., 2014 ; Meinecke et al., 2006). Bien que les systèmes de détection en 3D captent des images avec une résolution spatiale excellente, leur mise en place est onéreuse, lourde et requiert un espace considérable (Marcroft et al., 2015). De plus, l'utilisation des marqueurs collés au corps implique une manipulation peu tolérée par le nouveau-né prématuré (Figure 3 à droite). Dans ce cadre, les systèmes basés sur une détection en 2D représentent une alternative plus accessible par rapport à ceux basés sur une détection en 3D.

Les systèmes basés sur détection en 2D utilisent des caméscopes standards et même des dispositifs de smartphone (Kwong et al., 2019 ; Spittle et al., 2016). Ces systèmes présentent les avantages suivants : ils ne requièrent pas de marqueurs collés au corps ; ils sont moins coûteux ; ils occupent moins d'espace ; ils sont adéquats pour un usage dans un contexte naturel ; ils fournissent des images avec une haute résolution spatiale (Figure 4). Les systèmes basés sur détection en 2D ont permis l'identification des indicateurs d'anomalie suivants : une moyenne de mouvement petite et un écart type de mouvement petit (variabilité motrice réduite) ; une large variabilité du centre de gravité du mouvement (mouvement instable) (Adde et al., 2010, 2009).

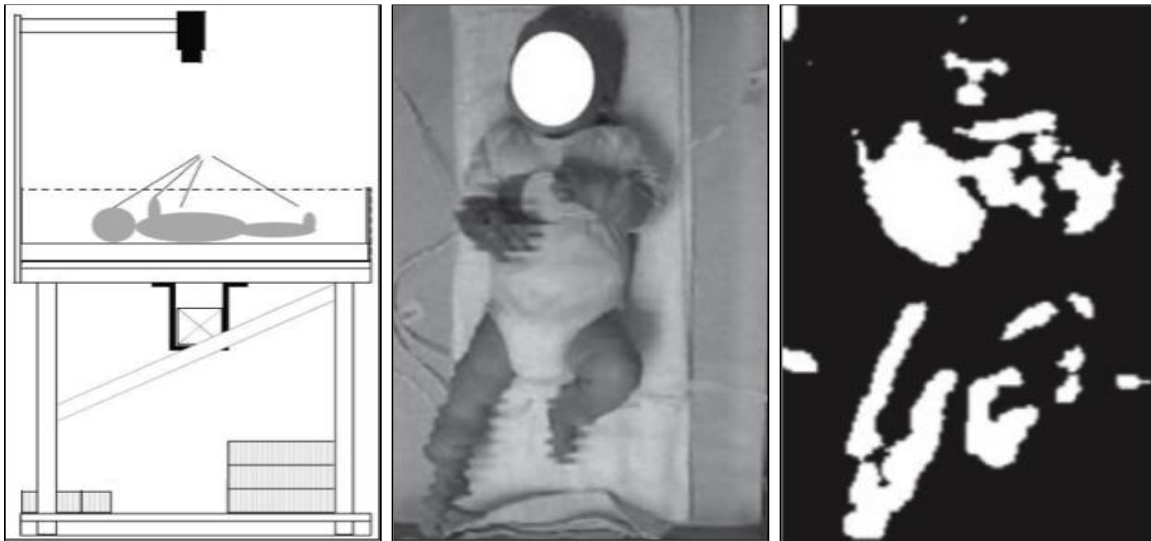
Figure 3. Système d'analyse des GMs sur détection en 3D



Équipement (à gauche) et positionnement des marqueurs (à droite). *Issu de Meinecke et al., (2006)*

Malgré ces avantages, les systèmes d'observation indirecte fournissent très peu d'information sur la vitesse et accélération des mouvements. Or, les GMs avec une vitesse réduite sont associés aux lésions cérébrales qui précèdent la PC (Hadders-Algra, 1993). En conséquence, les systèmes de détection indirecte actuels pourraient rencontrer des difficultés pour identifier des différences subtiles, mais cliniquement importantes, qui témoignent d'une anomalie.

Figure 4. Système d'analyse des GMs sur détection en 2D



Équipement de collecte de données (à gauche) ; vidéo d'entrée recadrée (au milieu) ; image animée dans l'outil d'analyse automatisée (à droite). Les pixels affichés en noir indiquent qu'aucun mouvement ne s'est produit entre les images, les pixels affichés en blanc représentent un mouvement. *Issu et traduit de Adde et al., (2010).*

3.3.2. Reconnaissance par détection directe

La détection directe se réalise par des accéléromètres attachés au corps qui captent la force inertielle générée par les membres en mouvement (Figure 5). Cela consiste à analyser des séries temporelles de données qui représentent l'accélération lors du mouvement. L'utilisation d'accéléromètres est répandue dans la recherche en vertu des avantages suivants : une haute résolution temporelle ; une accessibilité facile ; une petite taille des dispositifs à coller au corps. De cette manière, les détails et les différences subtiles du mouvement, qui s'avèrent fondamentales dans la clinique, sont mieux détectés. Ainsi, les indicateurs de risque détectés par des accéléromètres se caractérisent par : une corrélation entre l'accélération des mains et des pieds ; une même accélération observée sur tous les membres ; une accélération des pieds qui diffère de l'accélération moyenne globale ;

une accélération des mains en dehors de l'écart type de la moyenne globale ; des larges fluctuations des membres en mouvement (instabilité motrice) ; une grande imprévisibilité des mouvements des mains (degré de liberté important) ; la trajectoire d'un segment est similaire à celle d'autres segments (mouvements stéréotypés) ; une faible vitesse dans les articulations des jambes ; des mouvements périodiques des jambes (Pedalling) (Gravem et al., 2012; Heinze, Hesels, Breitbach-Faller, Schmitz-Rode, & Disselhorst-Klug, 2010; Karch et al., 2012 ; Ohgi, Morita, Loo, & Mizuike, 2008 ; Philippi et al., 2014). Cependant, en raison de la logistique de coller des capteurs au corps, l'applicabilité de cette technique se constate compliquée.

Figure 5. Accéléromètre et système de coordonnées



Accéléromètre (à gauche). Placement de l'accéléromètre (à droite). *Issu de Heinze, et al., (2010)*

3.3.3. Reconnaissance par des systèmes hybrides

Dans le but d'obtenir des informations temporo-spatiales précises, des systèmes hybrides, combinant la détection par vidéo avec celle par des accéléromètres, ont été développés (Marcroft et al., 2015). En effet, des systèmes hybrides ont permis

de préciser les critères suivants de risque : des séquences motrices produisant un signal d'onde courte, non fluide, irrégulière, sans apparence elliptique et non périodique ; une tendance à déplacer chaque membre d'une manière isolée et soudaine ; une corrélation croisée entre les membres en mouvements (Berge, Adde, Espinosa, & Stavdahl, 2008 ; Rahmanpour, 2009). Bien que les systèmes hybrides apportent une information plus détaillée des GMs, ces systèmes s'avèrent non adaptés au nouveau-né prématuré en raison des capteurs qu'il faut coller au corps.

3.3.4. Apprentissage sur des modèles cognitifs et sur des bases de données

L'analyse automatisée des GMs implique l'extraction des caractéristiques vidéographiques ou la description des signaux associés. Cela se base sur des modèles d'activité qui classent les GMs en typiques ou atypiques. D'une manière générale, les modèles d'activité se construisent en utilisant deux méthodes : l'apprentissage sur des modèles cognitifs ou l'apprentissage sur des bases de données cliniques (Chen et al., 2012). D'une part, l'apprentissage sur des modèles cognitifs exploite toutes les connaissances dans le domaine des GMs pour apprendre à une machine à les identifier. Ces connaissances sont modélisées avec des technologies d'ingénierie qui construisent des modèles logiques de la motricité précoce. Ces modèles analysent donc les GMs en utilisant un raisonnement formel (déductif ou inductif). Cette méthode d'apprentissage s'appelle l'approche cognitive ou top-down, car elle est axée sur un cadre conceptuel. D'autre part, l'apprentissage sur des bases de données cliniques utilise des cas de GMs pour entraîner une machine à les identifier. Cela recourt à des techniques d'exploration et d'apprentissage de données connues comme le « data-mining » et le « machine learning ». Il s'agit de trouver, auprès de cas de GMs, des règles pour identifier les anomalies selon un modèle probabiliste. Cela permet l'émergence d'une inférence qui consiste en une classification statistique connue comme approche bottom-up. D'ailleurs, chaque approche utilisée pour construire un modèle d'activité présente

des avantages et des inconvénients qui doivent être considérés lors du développement d'un système automatisé. Premièrement, les modèles d'activité basés sur l'approche cognitive sont clairs et faciles à utiliser. Cependant, ils présentent des difficultés pour gérer les informations temporelles et l'incertitude. Or, la précision et la stabilité d'un diagnostic, malgré la particularité du cas, sont fondamentales dans la clinique (Parniere, 2017). Deuxièmement, les modèles d'activité basés sur des données cliniques gèrent bien l'incertitude et l'information temporelle. Cependant, ils nécessitent un corpus clinique large pour bien apprendre le système à analyser la motricité, car ils sont moins précis lorsqu'ils sont construits sur des données qui ne sont pas représentatives de la population cible.

3.3.5. L'évaluation automatisée destinée au nouveau-né prématuré

Bien qu'aucun système d'analyse automatisé ne soit pas destiné au nouveau-né prématuré, il semblerait qu'un tel système doit inclure deux spécifications techniques : une détection indirecte par des images de vidéo en 2D ; une procédure d'analyse créée à partir d'une base de données sur les GMs d'un échantillon de nouveau-nés prématurés. Toutefois, la création d'une telle base de données s'avère compliquée, car les contraintes d'un milieu médicalisé et les risques de la prématurité rendent difficile l'accès aux enfants. Une telle base de données permettrait donc le développement d'un système automatisé d'évaluation neuromotrice adapté pour cette population.

3.3.6. Validité de l'évaluation automatisée

Les systèmes d'évaluation automatisée des GMs diffèrent quant aux résultats obtenus. Premièrement, les systèmes de détection par des accéléromètres présentent une sensibilité entre moyenne (0.72) et haute (1.00) pour identifier les

anomalies, alors qu'une spécificité entre basse (0.57) et haute (0.86) pour identifier les cas typiques. Deuxièmement, les systèmes de détection par vidéo démontrent une sensibilité haute (> 0.81) et une spécificité qui varie entre moyenne (0.70) et haute (0.88). Troisièmement, les systèmes hybrides présentent une sensibilité et une spécificité hautes (> 0.86 et > 0.90). Troisièmement, les systèmes hybrides présentent la meilleure performance. Néanmoins, ces conclusions ne peuvent pas être généralisées sur la population de nouveau-nés prématurés en raison de trois constats : aucun des systèmes publiés ne fut conçu pour évaluer la motricité du nouveau-né prématuré ; la validité rapportée correspond à des évaluations faites au 3e mois de vie ; la validation des systèmes automatisés considère des échantillons hétérogènes (Adde et al., 2010 ; Ihlen et al., 2019 ; Valle et al., 2015).

Encadré 3. Résumé du chapitre III

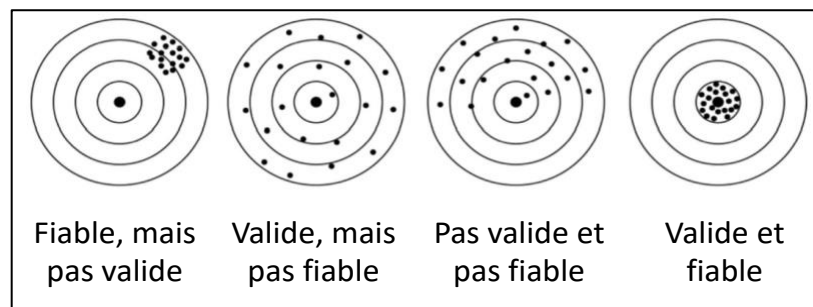
Les études sur la motricité spontanée permettent d'en tirer trois conclusions : le fœtus présente un riche répertoire moteur ; la plupart des patterns moteurs prénataux se retrouvent aussi chez le nouveau-né ; les GMs donnent un aperçu de l'intégrité du système nerveux en développement. Les GMs émergent durant la vie intra-utérine et présentent une continuité transnatale allant jusqu'au 6e mois de vie. Ces mouvements surgissent de l'activité intrinsèque des circuits sous-corticaux dénommés CPGs, et sont modulés par la sous-plaque corticale leur conférant leurs caractères distinctifs : la variabilité, la complexité et la fluidité. Chez l'enfant né prématurément, les mouvements atypiques s'associent notamment aux anomalies de la substance blanche traduites par des troubles cognitifs et moteurs. Malgré la pertinence de cette évaluation, son utilisation n'est pas répandue. Cela s'expliquerait par l'implication subjective de l'évaluateur pouvant fausser le diagnostic. Par conséquent, une approche qui associe l'informatique et la clinique conçoit une douzaine de systèmes automatisés capables d'évaluer objectivement les GMs. Aucun de ces systèmes n'est adapté ni validé pour le nouveau-né prématuré. Il semblerait que les meilleurs systèmes incluraient deux spécifications techniques : une détection par vidéo en 2D ; une analyse recourant aux bases de données sur les GMs. Cependant, les particularités des nouveau-nés prématurés rendent difficile l'acquisition d'une telle base de données, dont la création permettrait le développement d'un système automatisé, adapté et validé, pour cette population. Bien que l'évaluation automatisée des GMs constitue une approche prometteuse, elle doit relever deux défis : la conception et validation d'un outil automatisé pour le nouveau-né prématuré. Il s'avère donc que la validation est un impératif lors du développement de toute mesure. C'est pourquoi le chapitre suivant décrit les critères et la démarche de validation de n'importe quelle mesure neuromotrice.

IV. Validation des mesures neuromotrices

4.1. Propriétés psychométriques

La neurologie pédiatrique dispose d'une grande quantité de tests qui diffèrent quant aux caractéristiques et à la qualité (Majnemer & Snider, 2005). En effet, l'évaluation de la qualité d'une mesure repose sur des standards dénommés validité et fiabilité. La fiabilité et la validité sont des propriétés psychométriques qui jugent la convenance d'un test au moyen d'une démarche scientifique appelée validation. De ce fait, la validation est une approche transversale qui doit être adoptée dans toutes les étapes de développement d'un nouveau test (Mokkink, Terwee, Patrick, Alonso, Stratford, Knol, Bouter, & De Vet, 2010). Bien que la validité et la fiabilité diffèrent, elles tendent à se confondre. Ainsi, la validité estime la justesse d'un score et la fiabilité évalue sa stabilité. La figure 6 illustre cette différence.

Figure 6. Validité et fiabilité des mesures cliniques



Issu et traduit de de Souza, Alexandre, & Guirardello, (2017).

4.1.1. Validité

La validité détermine le degré auquel une mesure évalue vraiment le construit ou attribut qu'elle prétend évaluer. À part d'une révision conceptuelle exhaustive, la validité d'un test s'établit par la relation entre les scores que ce test produit et deux variables : les scores d'autres tests qui mesurent le même construit ; les événements cliniques qui affectent le construit que le test en question évalue (Aaronson & et al, 2002; Borsboom, Mellenbergh, & van Heerden, 2004; Cronbach & Meehl, 1955; Fermanian, 2005; Polit & Beck, 2006; Roach, 2006). C'est pourquoi la validité d'un test dépend de plusieurs facteurs : du type de construit mesuré, du contexte d'utilisation et de la population (Terwee et al., 2007).

4.1.2. Fiabilité

La fiabilité détermine la reproductibilité et consistance des scores produits par un test (Mokkink et al., 2010 ; Spittle et al., 2008). En termes pratiques, la fiabilité indique la concordance entre deux évaluations répétées d'un même test sur un même patient stable : une évaluation faite par un codeur à deux moments différents ; une évaluation faite par deux codeurs différents à un même moment (Kottner, Audig, Brorson, Donner, & Gajewski, 2011). Toutefois, la notion de « patient stable » dépend du construit⁴ évalué et de la population cible (de Vet, Terwee, Knol, & Bouter, 2006). Ainsi, la stabilité neuromotrice de l'enfant né prématurément est relative à son stade développemental et à sa situation clinique, ce qui, d'ailleurs, s'avère changeant. Par conséquent, la fiabilité d'un test se rapporte à la

⁴ « Construit ». La notion du construit est un terme générique issu de la psychométrie pour désigner tout sujet d'évaluation (ex. neuromotricité), directement mesurable ou non, bien défini et délimité dans un cadre théorique ou scientifique.

reproductibilité de ses résultats sur une population particulière dans un moment précis (Mokkink et al., 2012).

4.2. Indices statistiques de validation

La validation est une démarche empirique dont les résultats se présentent au moyen d'indices statistiques qui quantifient la validité et fiabilité d'une mesure (Aaronson & et al, 2002 ; Mokkink et al., 2010). Par conséquent, le choix d'un indice statistique dépend du « desing » de l'étude et de la nature des variables concernées (continues, ordinales et nominales). Ainsi, la validation des mesures neuromotrices destinées à l'enfant né prématurément utilise des indices épidémiologiques, d'association et d'accord.

4.2.1. Indices épidémiologiques

La sensibilité, la spécificité et les valeurs prédictives sont des indices les plus simples et les plus utilisés pour comparer deux tests dichotomiques (typique atypique) (Tableau 3). La sensibilité d'un test est la proportion de vrais positifs correctement identifiés, alors que la spécificité est la proportion de vrais négatifs correctement identifiés (Altman & Bland, 1994). Ainsi, la sensibilité indique la capacité d'une mesure à distinguer un trouble neuromoteur, alors que la spécificité se rapporte à la capacité pour distinguer la neuromotricité typique. Un test discriminant a une sensibilité et une spécificité proches de 100 % (1.00). Cependant, un test avec une sensibilité élevée peut avoir une spécificité faible, et inversement. Cela signifie qu'entre la sensibilité et la spécificité existe un compromis, et le clinicien doit prendre une décision en fonction de son importance relative. De même, la

sensibilité et la spécificité sont peu informatives sur la capacité diagnostique d'un test. C'est pourquoi il faut également calculer les valeurs prédictives. D'une part, la valeur prédictive positive (VPP) indique la probabilité que la maladie soit présente lorsque le test donne un résultat positif. D'autre part, la valeur prédictive négative (VPN) indique la probabilité que la maladie ne soit pas présente lorsque le test donne un résultat négatif. Cependant, ces indices doivent s'interpréter avec précaution, car la sensibilité et la spécificité sont influencées par les caractéristiques de l'échantillon et les indices prédictifs dépendent de la prévalence de la maladie.

Tableau 3. Tableau de décompte pour calculer des indices épidémiologiques

Dx	Non malade (GS = 0)	Malade (GS = 1)	Total
Négatif (Dx = 0)	VN = vrais négatifs	FN = faux négatifs	VN + FN = test négatif
Positif (Dx = 1)	FP = faux positifs	VP = vrais positifs	FP + VP = test positif
Total	VN + FP	FN + VP	VN + FP + FN + VP

GS, résultats du gold standard ou diagnostic d'intérêt ; Dx, diagnostic du test soumis à validation ; sensibilité = taux de vrais positifs = $VP/(FN+VP)$; spécificité = taux de vrais négatifs = $VN/(VN+FP)$; valeur prédictive positive = $VP/(FP+VP)$; valeur prédictive négative = $VN/(VN+FN)$.
Issu et adapté de (Eusebi, 2013).

4.2.2. Indices d'association

L'analyse de régression est une méthode utilisée dans la validation des mesures cliniques, car elle permet de modéliser l'association entre les scores d'un test et des données cliniques prédictives. L'analyse de régression linéaire permet de prédire les scores continus, alors que l'analyse de régression logistique permet de prédire les scores nominaux (Meurer & Tolles, 1986). Dans la régression linéaire, la variable

y (dépendante) est prédite à partir d'une variable x (indépendante) par l'équation suivante :

$$Y = b_0 + bX + E$$

Où le coefficient b_0 est l'interception (ordonnée à l'origine) et le coefficient bX est le gradient (Bewick, Cheek, & Ball, 2003). De son côté, la régression logistique prédit la probabilité que y se produise compte tenu des valeurs connues de x . Dans sa forme la plus simple, la régression logistique est représentée par l'équation suivante :

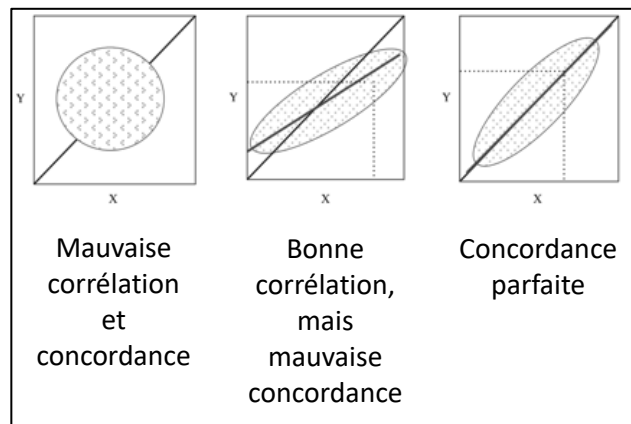
$$P(Y) = \frac{1}{1 + e^{-(b_0 + bx)}}$$

Où $P(Y)$ est la probabilité que y se produise, e est la base des logarithmes naturels, les coefficients b_0 et bX forment une relation linéaire (Field, Miles, & Field, 2012). Le coefficient de régression est un indice approprié pour quantifier la validité d'un test, car il permet de : décrire la relation entre les scores et les variables cliniques ; estimer la fréquence des scores ; pronostiquer les facteurs qui influencent les scores (Krzywinski & Altman, 2015 ; Schneider, Hommel, & Blettner, 2010). Plus particulièrement, l'analyse de régression univariée évalue l'association entre la variable dépendante et une variable indépendante, alors que l'analyse de régression multivariée évalue l'effet de plusieurs variables indépendantes sur la variable dépendante (Marill, 2004 ; Schneider et al., 2010).

4.2.3. Indices d'accord

Les indices d'accord quantifient la concordance entre des mesures répétées d'un même test (Figure 8). En effet, pour que deux mesures soient parfaitement concordantes, elles doivent être aussi bien corrélées (tous les points situés sur une droite quelconque) qu'identiques (tous les points situés sur la droite de pente 1 et d'ordonnée à l'origine 0). C'est pourquoi, au lieu d'utiliser des coefficients de corrélation ou d'association, la fiabilité de toute mesure doit être démontrée par des indices de concordance dont le choix dépend du type de score (continu, ordinal ou nominal) (Mandrekar, 2011 ; Watson & Petrie, 2010).

Figure 7. Représentation de la concordance entre deux mesures répétées



Issu de Elie & Colombet, (2011).

Le Kappa

Le coefficient Kappa (k) est utilisé pour évaluer la fiabilité entre des mesures répétées d'un test dont le score est nominal. Concernant les mesures neuromotrices destinées à l'enfant né prématurément, la fiabilité la plus simple et la plus courante

est celle de deux codages (ex. Codeur 1 et codeur 2) sur deux catégories (ex. Typique et atypique) (Bewick, Cheek, & Ball, 2004 ; Kraemer, Periyakoil, & Noda, 2002). En effet, le k est défini comme la proportion de jugements dans lesquels il y a un accord, tout en ajustant le degré d'accord que l'on peut s'attendre à ce que se produise par hasard (Chen, Faris, Hemmelgarn, Walker, & Quan, 2009). Pour ce faire, le k se base sur l'équation :

$$k = \frac{p_o - p_c}{1 - p_c}$$

Où p_o est la proportion d'accord observée et p_c est la proportion d'accord obtenue par chance. Par conséquent, le k définit la fiabilité par une valeur comprise entre 0 et 1 où une valeur proche de 1 indique un accord excellent entre les mesures comparées (Cohen, 1960). Cependant, il est nécessaire d'interpréter avec précaution le k , car il tend à baisser avec la prévalence de la maladie.

Le coefficient de corrélation intraclasse

Le coefficient de corrélation intraclasse (ICC) est utilisé pour examiner la fiabilité entre des mesures répétées d'un même test dont le score est continu. L'ICC montre la proportion de la variance totale qui peut s'attribuer aux différences entre les patients (Shrout & Fleiss, 1979). Bien qu'il existe plusieurs modèles d'ICC, la fiabilité des mesures neuromotrices pour l'enfant né prématurément peut utiliser un modèle mixte qui considère un effet « sujet » et un effet « méthode de mesure » aléatoire (Weir, 2005). Ce modèle permet de calculer la proportion de la variance totale (ρ) en utilisant l'équation suivante :

$$\rho = \frac{\textit{variance entrepatients}}{\textit{variance entrepatients} + \textit{variance entreméthodes} + \textit{variance d'erreur}}$$

L'ICC comprend des valeurs qui se situent entre 0 et 1 : celles qui se rapprochent de 1 suggèrent que la variance d'erreur est négligeable par rapport à la variance des patients ; celles qui se rapprochent de 0 dénotent que la variance d'erreur est extrêmement grande par rapport à la variance des patients (Koo & Li, 2016). Cependant, il est nécessaire d'interpréter avec précaution ces valeurs, car l'ICC tend à baisser dans un échantillon homogène.

4.2.4. Taille d'effet

Les études de validation utilisent des indices statistiques dont la significativité est établie avec la valeur de p . Bien que la valeur de p indique la force statistique d'une relation, elle n'indique pas forcément qu'une relation clinique importante ait été détectée (Stöckl, Rodríguez, & Thienpont, 2004). En effet, des différences sans intérêt réel peuvent être significatives sur de grands échantillons, alors que des effets cliniquement importants peuvent être non significatifs parce que l'échantillon est petit (Whitley & Ball, 2002). Or, de petits échantillons, sans différences statistiquement significatives, peuvent présenter des résultats cliniquement pertinents (Gardner & Altman, 1986). Par conséquent, la valeur de p n'indiquerait donc pas l'effet d'un facteur clinique d'intérêt (Fleiss, 1986). C'est pourquoi, dans un souci de pertinence clinique, les études qui valident une mesure neurologique pour le nouveau-né prématuré doivent rapporter des indices qui estiment sa pertinence clinique sur cette population. Il s'agit d'indices statistiques qui accompagnent la valeur de p pour juger l'effet sur la population, c'est-à-dire l'effet de taille. Les études qui valident les mesures neuromotrices pour l'enfant prématuré utilisent l'intervalle

de confiance (IC), le coefficient de Pearson (r) et l'odds ratio (OR) comme estimations de la population.

L'intervalle de confiance

L'IC (intervalle de confiance) est une plage de valeurs contenant la valeur vraie avec un risque d'erreur petit (Greenland et al., 2016). L'importance clinique de l'IC réside dans le fait qu'il montre les scores ou valeurs considérés comme plausibles pour la population, et ce à partir de l'observation d'un échantillon (McKenzie, Vida, Mackinnon, Onghena, & Clarke, 1997). Généralement, l'IC est fixé à 95 %, ce qui signifie que si une étude est répétée 100 fois, l'IC devrait inclure la valeur vraie à 95 reprises (Gardner & Altman, 1986). L'interprétation d'IC permet de généraliser les résultats : un petit IC indique que la valeur observée dans l'échantillon est proche de la valeur vraie dans la population ; un IC qui inclut la valeur de zéro indique que l'effet observé dans l'échantillon n'est pas significatif dans la population. Le r est un coefficient de corrélation qui estime la magnitude de l'effet dans une population donnée.

Le coefficient de Pearson

Le r (coefficient de Pearson) est un coefficient qui quantifie la corrélation entre deux variables continues. Le r est aussi utilisé pour évaluer l'effet sur la population par une valeur comprise entre 0 et 1 : une valeur de 0.1 indique un petit effet ; une valeur de 0.3 montre un effet moyen ; une valeur de 0.5 révèle un grand effet.

L'odds ratio

L'OR (odds ratio) indique le degré de dépendance entre deux variables dichotomiques (Simon, 2001). Il s'agit de comparer la cote d'un évènement (ex. Présence de signes neurologiques) qui arrive à un groupe A d'individus (ex. Enfants avec lésion cérébrale) avec la cote du même évènement qui arrive à un groupe B (ex. enfants n'ayant pas de lésion cérébrale). Étant donné que l'OR représente une comparaison par division, un ratio de 1.0 implique une similitude parfaite entre les deux groupes. Ainsi, plus l'OR s'écarte de 1.0, plus la différence entre les groupes est significative.

4.3. Aspects affectant la validation

La validation d'un test neuromoteur destiné à l'enfant né prématurément doit considérer les standards méthodologiques suivants : des mesures appropriées ; un échantillon représentatif ; des hypothèses formulées à priori ; des évaluations indépendantes et à l'aveugle ; un intervalle temporel adéquat entre les évaluations ; des indices statistiques appropriés (Bulpitt, 1987). À cela s'ajoute le contrôle du biais dans les domaines qui seront précisés dans les paragraphes suivants.

4.3.1. Échantillonnage

Un petit échantillon affecte la puissance des méthodes statistiques qui quantifient la validité des mesures utilisées. En effet, un petit échantillon risque de ne pas être représentatif de la population des enfants prématurés, car cela se rapporte à la difficulté de couvrir l'éventail des niveaux de risque clinique. De même, un petit

échantillon ($N \leq 30$) peut donner des résultats non significatifs, malgré un effet clinique important (Field, Miles, & Field, 2012). En outre, un recrutement en fonction de facteurs autres que les critères de recrutement restreint l'inclusion de sous-groupes pertinents. D'ailleurs, les résultats peuvent être biaisés en incluant des nouveau-nés à terme, puisque la motricité de ceux-ci diffère de celle des nouveau-nés prématurés. (Jeong et al., 2019 ; Wong et al., 2018 ; Albuquerque et al., 2017 ; Cardoso et al., 2017 ; Eeles et al., 2017 ; Ricci, Einspieler, & Craig, 2017 ; Gagnon, Cannon, & Weatherstone, 2016 ; Yeh et al., 2016 ; Morales, Bagur, Suc, Cazorla, & Girabent, 2015 ; Campbell & et al., 2013 ; Reuner et al., 2013 ; Heineman, Bos, & Hadders-Algra, 2008 ; Gagnon, Cannon, & Weatherstone, 2012 ; Heineman & Hadders-algra, 2011 ; Rose & Westcott, 2005). En somme, les problèmes d'échantillonnage affectent la généralisation des résultats, en rendant difficile de juger l'adéquation d'une mesure pour la population d'enfants prématurés rencontrés dans la pratique clinique.

4.3.2. Mesure

La comparaison entre des mesures qui diffèrent quant au composant moteur que chacune d'elles évalue rend difficile d'en tirer les conclusions sur leur validité. De même, les erreurs d'interprétation à cause d'une définition variable du seuil de coupure, d'une subjectivité importante dans le codage et d'une description inadéquate des items observés font augmenter le taux de diagnostics erronés. En outre, un test standardisé sur une population spécifique de nouveau-nés prématurés s'avère difficile à valider sur une population similaire, mais avec des origines ethniques différentes (Albuquerque et al., 2017 ; Gagnon, Cannon, & Weatherstone, 2016 ; Reuner, Fields, Wittke, Löprrich, & Pietz, 2013 ; Pin et al., 2010 ; Jeng et al., 2000 ; Seme-Ciglonecki, 2003 ; Campbell et al., 1995 ; Harris, Haley, Tada, & Swanson, 1984). En résumé, les problèmes liés à la mesure rendent difficile à estimer sa valeur clinique pour l'enfant prématuré.

4.3.3. Procédure d'évaluation

Le manque de contrôle de l'état d'éveil peut fausser les résultats de l'évaluation, puisqu'un état d'éveil ou de fatigue non approprié modifie la motricité des enfants. De même, la présence d'équipements d'aide à la ventilation et à l'alimentation, ainsi que les dispositifs de surveillance peuvent limiter l'évaluation neuromotrice du nouveau-né. En ce qui concerne le codage, le nombre restreint d'observateurs experts, la connaissance de l'histoire clinique des enfants, l'absence d'une norme d'OR appropriée et la non-indépendance des évaluations peuvent fausser les résultats. La dimension temporelle peut également représenter une source de biais dans la validation. Par conséquent, il faut tenir compte des aspects suivants : de l'effet des interventions précoces sur la motricité ; de l'intervalle de temps approprié entre chaque évaluation ; de faire passer les tests à des tranches d'âge gestationnel comparables (Lefebvre et al., 2016 ; Yeh et al., 2016a ; Gagnon et al., 2012 ; Bernhardt et al., 2011 ; Heineman, Middelburg, & Hadders-Algra, 2010 ; Heineman & Hadders-Algra, 2011 ; Jeng et al., 2000 ; Harris, Haley, Tada, & Swanson, 1984 ; Prechtl, 1974). En définitive, les difficultés dans la procédure peuvent fausser les résultats et biaiser la fiabilité d'un test pour l'enfant prématuré.

4.3.4. Traitement et rapport des données

L'utilisation d'indices statistiques non appropriés empêche d'apprécier la qualité des mesures évaluées. Particulièrement, l'utilisation de coefficients de corrélation pour évaluer la validité de critère ou la fiabilité s'avère non adaptée. Cela s'explique par le fait que les coefficients de corrélation indiquent l'association entre deux mesures, mais non la concordance entre elles. Un autre problème à éviter se rapporte aux données manquantes ou au taux d'attrition lors des évaluations longitudinales. Ces facteurs faussent les résultats, car les données manquantes

peuvent entraîner des différences significatives par rapport à l'échantillon original (Jeng et al., 2000 ; Albuquerque et al., 2017). En définitive, les difficultés liées au traitement et au rapport des données rendent difficile de juger de la validité et la fiabilité des mesures neuromotrices de l'enfant prématuré.

4.4. Utilité clinique

Bien que la validation s'avère une démarche essentielle pour juger la qualité des mesures cliniques, celle-là renseigne peu de leur utilité clinique. En effet, un test peut s'avérer fiable et valide, mais peu pratique d'utilisation dans un contexte clinique. Un test est d'autant moins utile cliniquement qu'il prend beaucoup de temps, qu'il nécessite des ressources excessives pour pouvoir l'appliquer ou qu'il requiert une formation pour le codage. Par conséquent, dans un souci d'utilité clinique accrue, plus le temps, les exigences et les manœuvres pour faire passer un test seront diminués, plus ce test sera avantageux. En d'autres mots, bien que l'utilité clinique ne fasse pas l'objet de validation, car elle n'est pas une propriété psychométrique, elle est pourtant un paramètre fondamental pour juger de la convenance d'une mesure (Jeong et al., 2019). Cela est d'autant plus important en ce qui concerne les mesures neuromotrices des nouveau-nés prématurés, car elles doivent répondre aux défis imposés par les particularités de cette population. Plus spécialement, l'utilité clinique d'une telle mesure, qu'elle soit qualitative, quantitative ou automatique, se rapporte aux paramètres suivants : diminuer considérablement le temps d'évaluation ; ne pas avoir besoin d'une formation particulière ; faciliter la manutention du matériel ; proposer un codage facile, non susceptible d'interprétations ambiguës (Keszei, Novak, & Streiner, 2010).

Encadré 4. Résumé du chapitre IV

L'évaluation est essentielle dans la pratique clinique. Elle se base sur le diagnostic et le pronostic obtenus par des mesures qui doivent présenter une bonne qualité. Au contraire, il existerait un risque de résultats biaisés qui conduisent à des conclusions erronées. Parmi les paramètres de qualité de toute mesure neurologique de l'enfant prématuré se compte la validité qui juge la justesse des scores et la fiabilité qui juge la stabilité des scores. La validité et la fiabilité d'une mesure sont évaluées au moyen des études de validation. Les études de validation constituent une démarche expérimentale permettant de tester la pertinence d'une mesure pour être appliquée sur une population spécifique à un moment précis. Cela est particulièrement important pour les mesures neuromotrices destinées à l'enfant prématuré, car sa motricité présente les particularités suivantes : elle diffère de celle de l'enfant né au terme ; elle change en fonction de facteurs cliniques, maturationnels et développementaux ; elle ne reflète pas toujours la fonction neurologique. Avant d'utiliser une mesure neuromotrice chez les enfants prématurés, sa qualité doit être démontrée par des preuves empiriques de haute qualité qui attestent les aspects suivants : l'exhaustivité des items pour évaluer la neuromotricité ; la concordance des scores avec d'autres mesures neurologiques ; le changement des scores en fonction de complications cliniques qui affectent la fonction neurologique néonatale ; la reproductibilité des scores. Une mesure neuromotrice pour le nouveau-né prématuré doit aussi avoir une utilité clinique importante. L'utilité clinique d'une telle mesure se traduit en quatre avantages : facilité pour l'appliquer ; diminution du temps d'expérimentation ; réduction du degré de manipulation de l'enfant ; suppression d'une formation complémentaire ; facilité pour acquérir et installer le matériel. Ces qualités, associées à une bonne validité et une bonne fiabilité, sont fondamentales pour le développement de la meilleure mesure clinique destinée à l'évaluation neuromotrice chez le nouveau-né prématuré.

V. Bilan théorique, problématique et objectifs

Les enfants qui naissent prématurément (avant 37 SA) courent le risque de développer des troubles neurologiques de type sensoriel, cognitif et moteur. Ce risque est exacerbé par un petit âge gestationnel et des complications néonatales telles que la DBP (dysplasie bronchopulmonaire), l'EUNC (l'entérocolite ulcéro-nécrosante), les infections et les lésions cérébrales. De même, les interventions développementales pour l'enfant né prématurément sont plus bénéfiques lorsqu'elles commencent à la période néonatale, quand l'enfant est encore à l'hôpital et le développement cérébral est en pleine expansion. Bien que toute intervention dépende de l'identification des nouveau-nés prématurés plus à risque, le manque d'une mesure neuromotrice objective demeure une question pertinente. Cela peut s'expliquer par le fait qu'une telle mesure doit répondre aux particularités de cette population et remplir des exigences de validation qui s'avèrent être compliquées à réaliser dans un contexte médicalisé.

En effet, le TIMP (Test of Infant Motor Performance) et l'analyse des GMs sont les mesures neuromotrices les plus adaptées aux particularités du nouveau-né prématuré. D'une part, elles coïncident sur le fait qu'elles considèrent l'analyse de la motricité spontanée, pourtant leurs propriétés psychométriques restent encore à établir. D'autre part, elles diffèrent quant au degré de manipulation de l'enfant. Bien que les GMs n'impliquent aucune manipulation du nouveau-né prématuré, cette méthode n'est pas utilisée dans les services de néonatalogie à cause des contraintes liées à la mesure et la procédure. Pour pallier ces inconvénients, des systèmes automatisés qui évaluent d'une manière objective les GMs ont été développés. Cependant, aucun de ces systèmes n'est adapté au nouveau-né prématuré qui est hospitalisé. Il semblerait donc qu'un système adapté à cette

population doit comporter deux caractéristiques techniques. Premièrement, une analyse automatisée d'images de vidéo en 2D. Deuxièmement, un modèle d'activité créé à partir d'une base de données sur les GMs.

5.1. Problématique

Cette thèse porte sur la constitution et validation d'une base de données sur les GMs auprès des nouveau-nés prématurés qui sont hospitalisés. Cette base de données, en tant que mesure neuromotrice, permettra le développement d'un système automatisé destiné au nouveau-né prématuré. Pour ce faire, nous présentons deux études qui répondent à deux objectifs généraux.

5.2. Objectifs

- 1) Décrire longitudinalement les GMs d'un échantillon d'enfants nés prématurément.
- 2) Évaluer l'adéquation des GMs pour mesurer la neuromotricité d'un échantillon d'enfants nés prématurément.

Encadré 5. Résumé du chapitre V

L'identification précoce des enfants prématurés à plus de risque neurologique requiert d'un outil objectif et adapté aux nouveau-nés prématurés hospitalisés. Le développement d'un tel outil requiert d'une mesure neuromotrice qui permet la construction d'un modèle d'activité pour le programmer. C'est pourquoi le travail de cette thèse concerne la création d'une mesure neuromotrice qui consiste en une base de données des GMs auprès des enfants nés prématurément. Pour ce faire, cette thèse comporte deux études empiriques. La première étude concerne la constitution de la mesure en question par l'enregistrement, l'évaluation longitudinale et la description des GMs d'un échantillon d'enfants nés prématurément. La deuxième étude vise la validation de la mesure qui sera constituée dans la première étude, afin d'évaluer son adéquation particulière au nouveau-né prématuré qui est à l'hôpital. Ces deux démarches empiriques seront détaillées dans les chapitres VI et VII.

VI. Constitution d'une base de données des mouvements généraux

Les mesures neuromotrices destinées au nouveau-né prématuré doivent : éviter toute stimulation de l'enfant ; refléter des changements développementaux ; représenter l'évolution clinique de la neuromotricité ; identifier les enfants ayant des anomalies neuromotrices et ceux n'en ayant pas (Sweeney & Blackburn, 2013; Blauw-Hospers & Hadders-Algra, 2005; Back, Riddle, & McClure, 2007; Hadders-Algra, 2014; Gosselin, Gahagan, & Amiel-Tison, 2005; Hadders-Algra, 2014). Or, des études suggèrent le besoin de ce type de mesures, car il n'existe pas de test considéré comme le *gold standard* (mesure de référence) pour évaluer la neuromotricité chez cette population (Brown & Spittle, 2014; Majnemer & Mazer, 1998). C'est pourquoi cette première étude porte sur la constitution d'une base de données de GMs auprès de nouveau-nés prématurés, et cela pour développer un outil d'évaluation neuromotrice basée sur vidéographie en 2D. Par conséquent, nous émettons trois hypothèses qui décrivent les GMs en tant que mesure neuromotrice : les GMs détectent les changements neuromoteurs qui surviennent au cours du développement ; les GMs permettent de tracer la trajectoire clinique de la neuromotricité ; les GMs permettent de distinguer entre les enfants ayant des anomalies neuromotrices de ceux qui n'en ont pas.

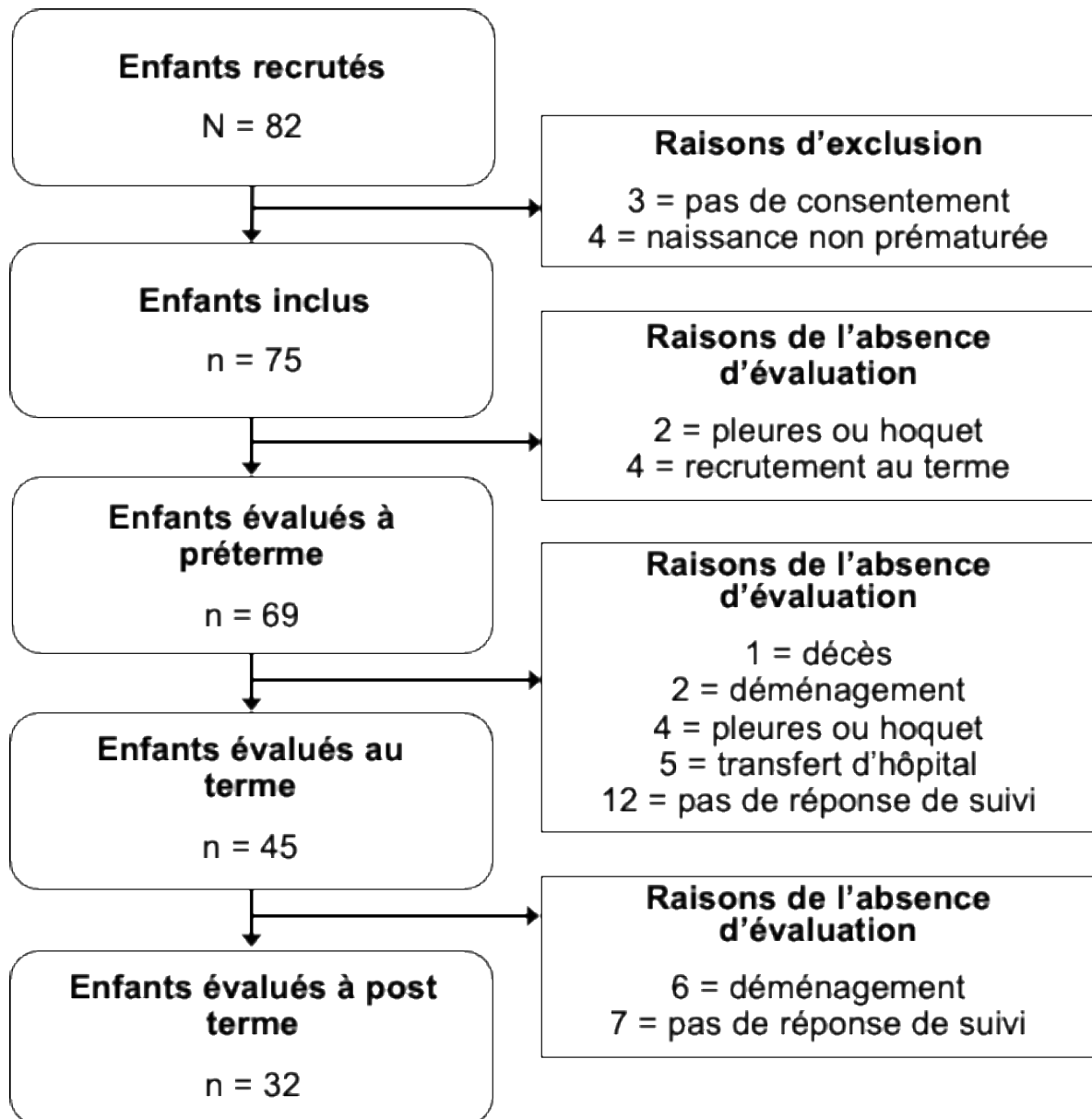
6.1. Méthodologie

Cette étude a obtenu un avis favorable du Comité de Protection de Personnes d'Île-de-France (Idf 5 : d-10-16). Elle est prospective avec un suivi longitudinal d'une cohorte et est de nature observationnelle auprès de nouveau-nés prématurés.

6.1.1. *Participants*

La figure 8 décrit le processus de recrutement et évaluation des GMs. Entre décembre 2016 et décembre 2017, les participants furent recrutés au sein du service de néonatalogie de deux hôpitaux universitaires à Paris : Armand-Trousseau (n = 33) et Necker-Enfants Malades (n = 49). Quatre-vingt-deux enfants ont été recrutés dont 7 enfants ont été exclus, car ils présentaient au moins un critère de non-inclusion. L'échantillon final est composé de 75 enfants avec une moyenne d'âge gestationnelle de 31.6 SA (SD = 2.3), ainsi qu'une moyenne de poids de naissance de 1534.3 g (SD = 413.8). Les parents des enfants étaient contactés en fonction de l'admission à l'hôpital. Pour être inclus, les enfants devaient être nés entre moins de 28 SA et 36.6 SA et les parents devaient avoir signé le formulaire de consentement. Tous les niveaux de la prématurité ont été considérés. Les critères d'exclusions étaient la minorité des parents, la non-maîtrise de la langue française, ou le refus du pédiatre. Les enfants bénéficiant d'un traitement sédatif ont également été exclus.

Figure 8. Recrutement et évaluation des GMs



6.1.2. Recueil des données

La motricité spontanée des participants fut enregistrée aux périodes préterme, terme et post terme pour apprécier les changements développementaux et tracer l'évolution clinique des GMs. L'enregistrement vidéo durait 10 minutes et était réalisé au moyen d'une caméra GoPro qui, télécommandée à la distance, capturait la motricité de l'enfant en 2D. La caméra, fixée sur un trépied, était positionnée à un mètre au-dessus du matelas, le focus sur le centre de gravité de l'enfant. Chaque enfant était filmé après les soins et le repas à la suite d'un cycle de sommeil. Pour les enfants nourris par sonde, en continu, l'enregistrement avait lieu après les soins et pendant le fonctionnement de la sonde. En couche et/ou en body, l'enfant était placé au sein de sa couveuse ou de son lit en position décubitus dorsal et symétrique, avec la tête positionnée sur la ligne médiane. Les équipements médicaux ou autres objets limitant le mouvement des enfants étaient retirés au maximum. Les adultes ne devaient pas interagir avec l'enfant lors de l'enregistrement vidéo. L'état d'éveil de l'enfant était contrôlé par l'observation comportementale avant et tout au long de l'enregistrement (Tableau 4). Cela garantissait l'acquisition des GMs dans un état optimal (état IV) où la motricité spontanée s'exprime le mieux (Prechtl, 1974). Si l'enfant n'était pas dans l'état optimal, l'enregistrement était arrêté et reporté à un moment ultérieur, afin que l'évaluation neuromotrice n'aboutisse pas à un diagnostic erroné (Gagnon et al., 2012; Prechtl, 1974).

Un montage vidéo permet de ne sélectionner que les séquences de GMs à partir des enregistrements vidéo de la motricité spontanée. Pour être sélectionnée, chaque séquence devait avoir une durée minimale de 10 secondes et correspondre à la définition des GMs. Les GMs sont des mouvements auto générés, qui ne peuvent pas être liés à des stimuli externes, où toutes les parties du corps participent à une même séquence qui présente un commencement et une fin définis (Prechtl, 1990; Prechtl, 2001). Toutes les périodes de hoquet ou de pleurs étaient supprimées, afin d'éviter le risque d'un faux diagnostic (Einspieler et al., 2004). Les

séquences de GMs interrompues par une stimulation externe à l'enfant ou des mouvements autour de lui ont également été supprimées. Il en a été de même pour les séquences de GMs présentant des problèmes techniques : éclairage inadéquat, changements brusques de luminosité, tremblements de l'image ou déplacements de la caméra. Au cours des trois périodes d'évaluation confondues, 146 vidéos de la motricité spontanée ont été obtenus : 69 à préterme, 45 à terme, 32 à post terme. Au total, 785 séquences de GMs furent sélectionnées : 347 à préterme, 257 au terme, 181 à post terme. La moyenne d'âge corrigée au moment de l'évaluation des GMs est de 35.0 AGc (SD = 1.3) à préterme, 38.6 AGc (SD = 1.0) au terme, 13.5 SPT (SD = 2.9) à post terme.

Tableau 4. Les états comportementaux du nouveau-né

	Définition	Description comportementale
I	Sommeil calme	Yeux fermés, sursauts, pas de mouvements des membres, possibles mouvements rythmiques de la bouche, respiration calme et régulière, fréquence cardiaque stable.
II	Sommeil agité	Yeux fermés, mouvements oculaires, sursauts, nombreux mouvements des membres, sourires et grimaces, contractions, respiration et rythme cardiaque irréguliers.
III	Veille calme	Yeux ouverts, peu ou pas de mouvements, respiration calme et régulière, fréquence cardiaque stable.
IV	Veille active	Yeux ouverts, sourires et grimaces, mouvements fréquents des membres, contractions, respiration et rythme cardiaque irrégulier.
V	Pleurs ou cris	Yeux ouverts ou fermés, vocalisations, activité motrice intense et désordonnée, contractions, respiration et rythme cardiaque irrégulier.

Issu et adapté de Prechtl (1974)

6.1.3. Évaluation des mouvements généraux

Les GMs ont été évalués d'une manière globale et détaillée à partir des séquences des GMs (n = 785) étant sélectionnées au préalable. Selon l'approche qualitative, l'évaluation globale des GMs se base sur l'analyse d'une *gestalt* pour coder nominativement (typique/atypique) la variabilité, complexité et fluidité des GMs. La variabilité des GMs se définit par les changements de vitesse et l'occupation de l'espace, voire les changements de direction par les membres en mouvement. La complexité s'établit si tous les membres et toutes les articulations participent à la séquence des GMs. La fluidité implique la fluctuation dans le commencement et fin des GMs, sans changements brusques. Les GMs atypiques se caractérisent donc par une réduction de la variabilité, de la complexité et de la fluidité.

Selon l'approche quantitative, l'évaluation détaillée à préterme et terme se base sur une feuille de codage divisée en deux parties, dont le score est représenté par un nombre (Annexe 1). La première partie code les GMs en fonction des classifications : typique, si le mouvement est variable, complexe et fluide ; pauvre répertoire, si le mouvement est monotone avec une occupation limitée de l'espace et une vitesse lente ; chaotique, si le mouvement est désorganisé, d'une grande amplitude et la vitesse exagérément rapide ; crampe synchronized, si les membres se contractent et se relaxent simultanément avec rigidité ; Hypokinétique, si aucun mouvement n'est observé. Dans le cas d'une classification hypokinétique, l'évaluation détaillée n'est pas réalisée. La séquence des GMs est codée avec trois sous-scores : « 2 », si elle est variable ; « 1 », si elle est monotone ou cassée ; « 0 », si elle est synchronisée ou désorganisée. La deuxième partie de l'évaluation se concentre sur le cou et le tronc d'une part et les membres inférieurs et les membres supérieurs d'autre part. Pour chacun des items, une description de 20 critères optimaux est fournie et codée avec trois sous-scores : « 2 », si le critère est satisfait ; « 1 », si le critère est occasionnellement satisfait ; « 0 », si le critère n'est pas satisfait. La somme des sous-scores produit le score total allant de la valeur minimale de « 5 » pour la performance la moins bonne à la valeur maximale de « 42 »

pour une performance optimale. Concernant l'évaluation de la période post terme, le score détaillé de la grille de codage tient compte de 5 critères : la qualité des GMs fidgety ; la présence d'autres mouvements accompagnant les GMs (ex. Kicking) ; la qualité d'autres mouvements accompagnant les GMs ; la qualité de la posture ; le caractère des mouvements (Annexe 2). Le score d'un critère satisfait peut-être de « 12 » (pour les GMs fidgety typiques) ou « 4 » (pour les autres mouvements). Un critère occasionnellement satisfait est codé de « 2 ». Un critère non acquis est noté de « 1 ». La somme de ces sous-scores produit le score total pouvant aller du score minimal de « 5 » indiquant la performance la moins bonne au score maximal de « 28 » indiquant une performance optimale.

L'évaluation globale permet d'obtenir le diagnostic des GMs, alors que l'évaluation détaillée permet d'obtenir le score optimal des GMs. Ces deux évaluations furent réalisées par AV sans connaître la situation clinique des enfants. Pour cette première étude, JP a validé le codage d'AV sur 37 % des GMs dans les trois périodes confondues. AV et JP ont la certification avancée pour l'évaluation des GMs. Dans le cadre d'un désaccord, le cas en question faisait l'objet de discussion jusqu'à ce que l'accord fût trouvé. En cas de doute sur un cas, l'avis d'une tierce personne était demandé. Il s'agissait de la docteure Christa Einspieler.

6.1.4. Traitement et analyse des données

Trois mesures des GMs ont été utilisées. Premièrement, le diagnostic dichotomisé des GMs comme typiques et atypiques en utilisant les classifications observées dans les trois périodes d'évaluation : durant la période préterme et terme (typique et atypique : pauvre répertoire, cramped synchronized ou chaotique) ; durant la période post terme (typique et atypique : fidgety atypique ou fidgety intermittent). Deuxièmement, le score optimal des GMs obtenu lors de l'évaluation détaillée des GMs. Troisièmement, les sous-scores des critères optimaux des GMs (20 à préterme et terme ; 5 à post terme) dont la somme produit le score optimal.

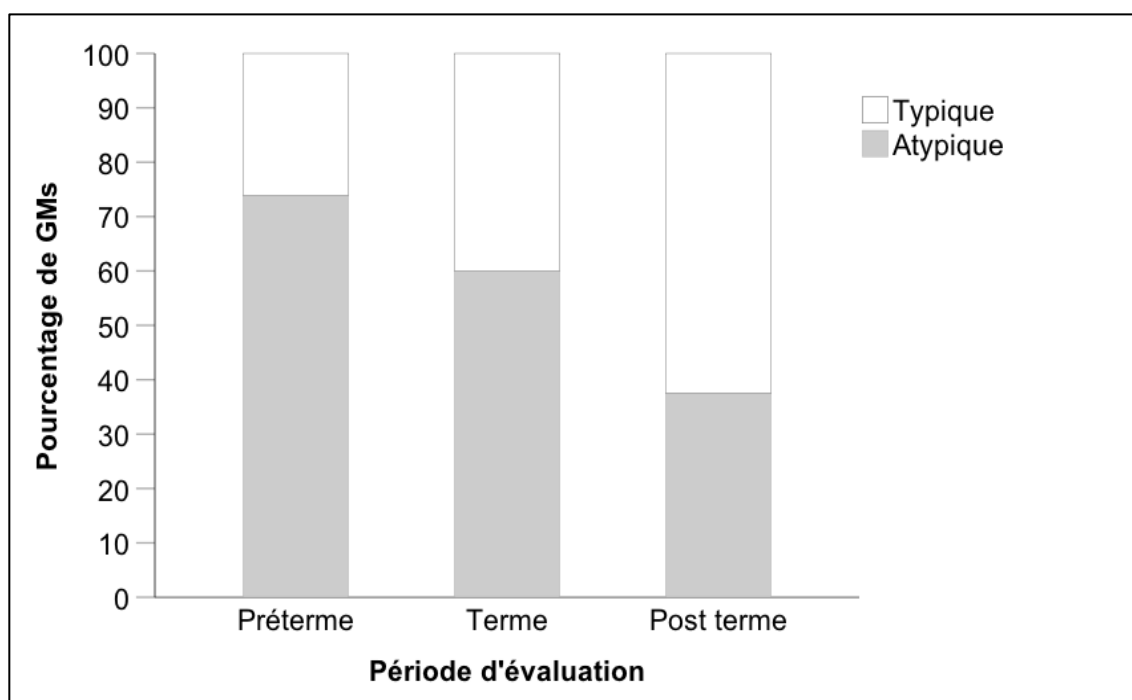
Les analyses statistiques furent réalisées sur chaque enfant ayant été évalué au moins une fois au cours des trois périodes : préterme, terme et post terme. En ce qui concerne le diagnostic (typique/atypique) des GMs, les trois périodes d'évaluation ont été comparées entre elles par des analyses intra-individuelles mesurant les changements développementaux. L'analyse intra-individuelle, pour comparer le score optimal des GMs dans les trois périodes, n'a pas été réalisée, car la feuille de codage utilisée à préterme et terme diffère de celle utilisée à post terme. Des analyses interindividuelles, sur le diagnostic et le score optimal des GMs, furent réalisées pour évaluer les différences cliniques à chaque période.

Les données nominales se présentent en termes de fréquence et proportion (ex. pourcentage du diagnostic des GMs). La médiane (Méd) et des boîtes aux moustaches sont rapportées pour les variables quantitatives ne suivant pas la loi d'une distribution normale. Le test Q de Cochran fut utilisé pour comparer les variables dichotomiques et appariées (ex. comparaison entre les trois périodes d'évaluation concernant le diagnostic des GMs). Le test McNemar ou le k (Kappa) fut utilisé pour évaluer l'association entre les variables catégorielles et appariées (ex. comparaison des classifications des GMs à travers les trois périodes d'évaluation). Des tests non paramétriques furent utilisés pour les variables quantitatives dont la distribution ne suit pas la loi normale. Ainsi, le test Z de Kolmogorov-Smirnov et J de Jonckheere furent utilisés pour les variables quantitatives non appariées (ex. comparaison du score optimal des GMs typiques à celui des GMs atypiques). Le test T de Wilcoxon signed-rank fut utilisé pour les variables quantitatives appariées (ex. comparaison du score optimal des GM à préterme à celui des GMs à terme). L'analyse de régression linéaire fut utilisée pour prédire les variables continues (ex. prédiction du score optimal des GMs à partir du diagnostic dichotomisé des GMs). La valeur de $p < .05$ était considérée statistiquement significative et accompagnée d'un IC BCa (intervalle de confiance sur la méthode bootstrap pour corriger le biais) fixé à 95 %. La taille de l'effet est rapportée par l'OR (odds ratio) ou le r (coefficient de Pearson). Les analyses statistiques ont été réalisées avec IBM SPSS Statistics.

6.2. Résultats

La figure 9 illustre la distribution du diagnostic (typique/atypique) des GMs au cours des trois périodes d'évaluation. La proportion des GMs atypiques diminue au cours des périodes d'évaluation : 74 % (n = 51) à préterme, 60 % (n = 27) au terme et 37 % (n = 12) à post terme. Le diagnostic des GMs pour chaque enfant variait selon la période d'évaluation. De ce fait, au cours des périodes d'évaluation, 47 % (n = 39) des enfants ont obtenu au moins un diagnostic de GMs typiques et 45 % (n = 37) n'ont obtenu que le diagnostic de GMs atypiques. Plus particulièrement, 5 % (n = 4) des enfants avaient des GMs typiques à préterme suivis des GMs atypiques à terme ou post terme et 20 % (n = 24) avaient des GMs atypiques à préterme suivis des GMs typiques à terme ou post terme. Le diagnostic des GMs à travers les trois périodes ne diffère pas (Cochran) $X^2(2) = 4.667$, $p = NS$.

Figure 9. Diagnostic des GMs



Le tableau 5 synthétise la relation entre le diagnostic (typique/atypique) des GMs à préterme et celui des GMs à terme. Il existe une relation significative entre le diagnostic des GMs à préterme et celui à terme, (McNemar) $X^2(1, n = 37) = 4.083$, 95 % [0.97, 33.6], $p < .05$, OR (odds ratio) = 5.7. Ainsi, les GMs atypiques à préterme entraînent plus de GMs atypiques à terme (90 %) que les GMs typiques à préterme. Basé sur l'OR, la chance d'obtenir des GMs atypiques à terme est 5.7 fois plus élevée s'ils étaient atypiques à préterme que s'ils étaient typiques à préterme.

Tableau 5. Diagnostic des GMs à préterme et terme

Préterme	Terme		Total
	Atypique	Typique	
Atypique	19	10	29
Typique	2	6	8
Total	21	16	37

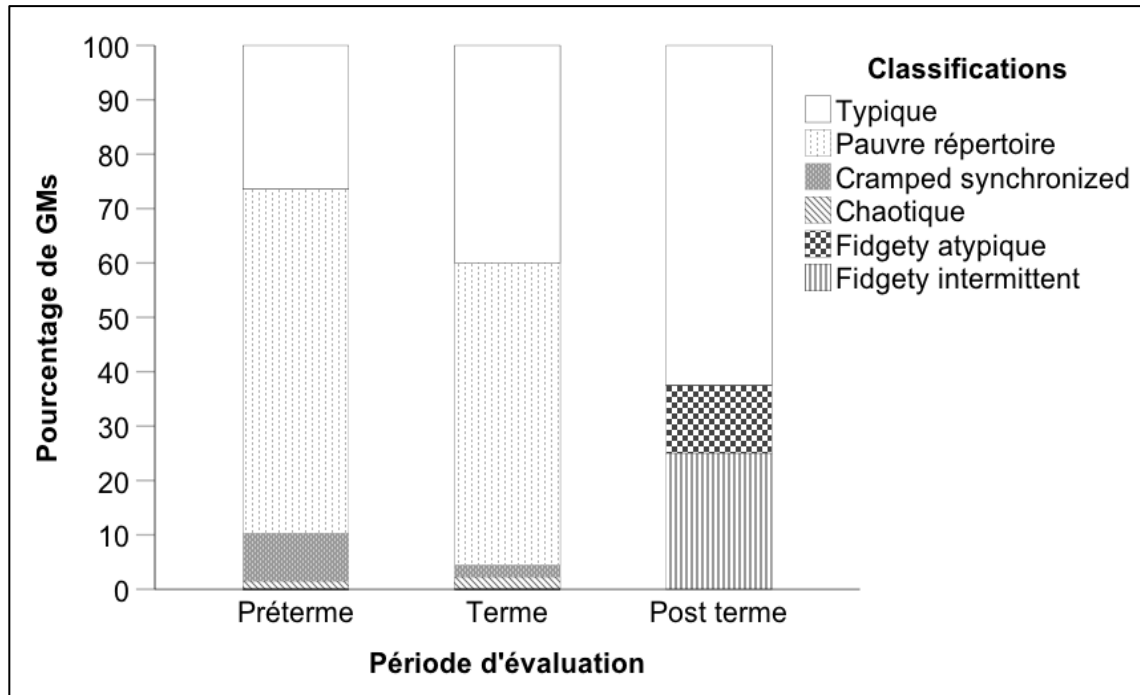
Le tableau 6 montre la relation entre le diagnostic (typique/atypique) des GMs à préterme et celui à post terme. Le diagnostic des GMs à préterme et celui des GMs à post terme n'ont pas de relation significative, $X^2(1, n = 29) = 3.765$, 95 % [14, 3.35], $p = NS$, OR = .67. Autrement dit, un diagnostic atypique à préterme ne prédit pas un diagnostic atypique à post terme.

Tableau 6. Diagnostic des GMs à préterme et post terme

Préterme	Post terme		Total
	Atypique	Typique	
Atypique	7	13	20
Typique	4	5	9
Total	10	19	29

La figure 10 montre la distribution des classifications des GMs dans les trois périodes d'évaluation. La classification pauvre répertoire est la plus fréquemment observée à préterme (43 soit 62 % des GMs) et également à terme (25 soit 55 %). La classification typique est la plus fréquemment observée (20 soit 62 % des GMs) à post terme. Au cours des trois périodes, 8 % (n = 7) des enfants ont obtenu au moins une classification de GMs cramped synchronized. Le 4 % (n = 3) avait des GMs cramped synchronized à préterme suivis des GMs pauvre répertoire ou cramped synchronized à terme, alors que 2% (n = 2) avait des GMs cramped synchronized à préterme suivis des GMs typiques à post terme. Aucun enfant n'a régulièrement présenté de GMs cramped synchronized. Le 8 % (n = 7) des enfants présentent des GMs pauvre répertoire ou chaotique à préterme qui se transforment en des GMs fidgety atypiques ou intermittents à post terme.

Figure 10. Pourcentage de classifications des GMs dans les trois périodes



Le tableau 7 synthétise la relation entre les classifications des GMs à préterme et celles des GMs à terme. Les classifications des GMs à préterme et celles des GMs à terme ne présentent pas de relation significative, (McNemar), $X^2(3, n = 37) = 6.455$, $p = NS$.

Tableau 7. Classifications des GMs à préterme et terme

Préterme	Terme				Total
	Typique	Pauvre répertoire	Cramped synchronized	Chaotique	
Typique	6	2	0	0	8
Pauvre répertoire	9	15	0	0	24
Cramped synchronized	1	2	1	0	4
Chaotique	0	0	0	1	1
Total	16	19	1	1	37

Le tableau 8 montre la relation entre les classifications des GMs à préterme et celles des GMs à post terme. Les classifications des GMs à préterme et celles des GMs à post terme ne présentent pas de relation statistiquement significative, (Kappa) $k = -.02$, $p = NS$.

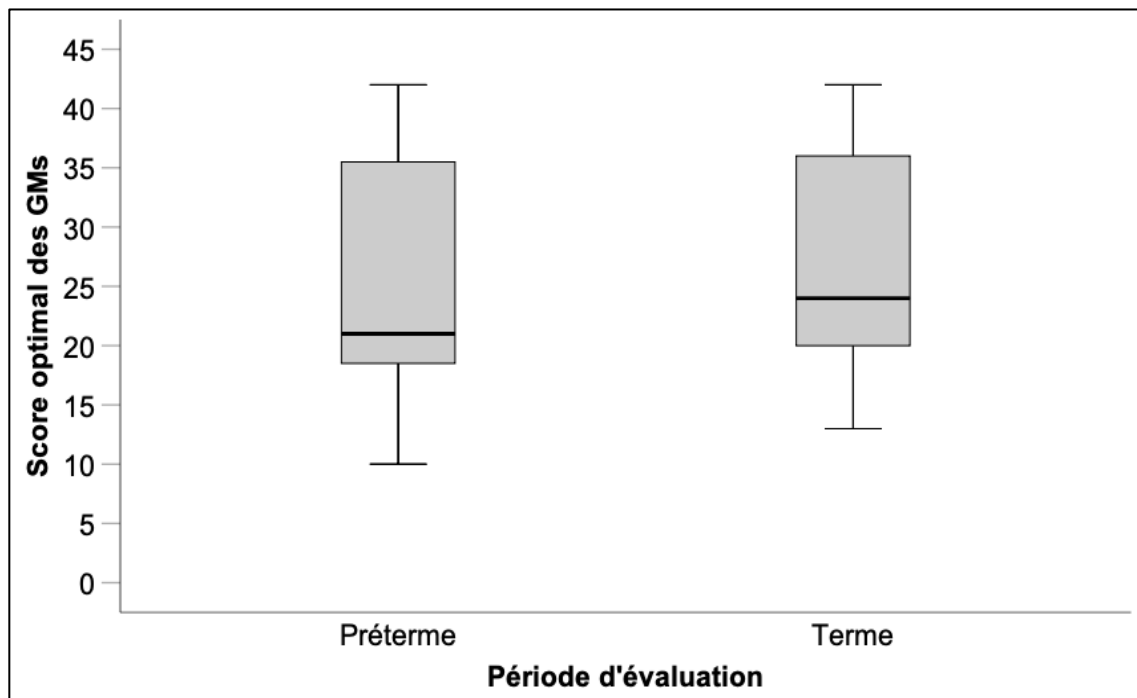
Tableau 8. Classifications des GMs à préterme et post-terme

Préterme	Postterme			Total
	Typique	Fidgety atypique	Fidegty intermittent	
Typique	5	2	2	9
Pauvre répertoire	11	2	4	17
Cramped synchronized	2	0	0	2
Chaotique	0	0	1	1
Total	18	4	7	29

La figure 11 compare la distribution du score optimal des GMs à préterme à celle du score optimal des GMs à terme. Les scores des GMs à terme (Méd = 24) sont

meilleurs que ceux des GMs à préterme (Méd = 21), (Wilcoxon signed-rank) $T = 427.000$, $p < .05$, $r = .43$. L'analyse de régression linéaire montre que les scores des GMs à préterme prédisent significativement ceux des GMs à terme, $F(1,35) = 7.179$, $p < .05$. Ainsi, plus les scores sont élevés à préterme, plus les scores seront élevés à terme. La valeur de R^2 indique que les scores à préterme expliquent 17 % de la variation des scores à terme. De même, le diagnostic (typique/atypique) des GMs à préterme prédit le score optimal des GMs à terme, $F(1,35) = 6.191$, $p < .05$. Ainsi, les GMs atypiques de la période préterme présenteront un score 9 fois plus bas à terme que les GMs typiques de la période préterme. La valeur de R^2 montre que le diagnostic des GMs à préterme explique 15 % de la variation des scores à terme.

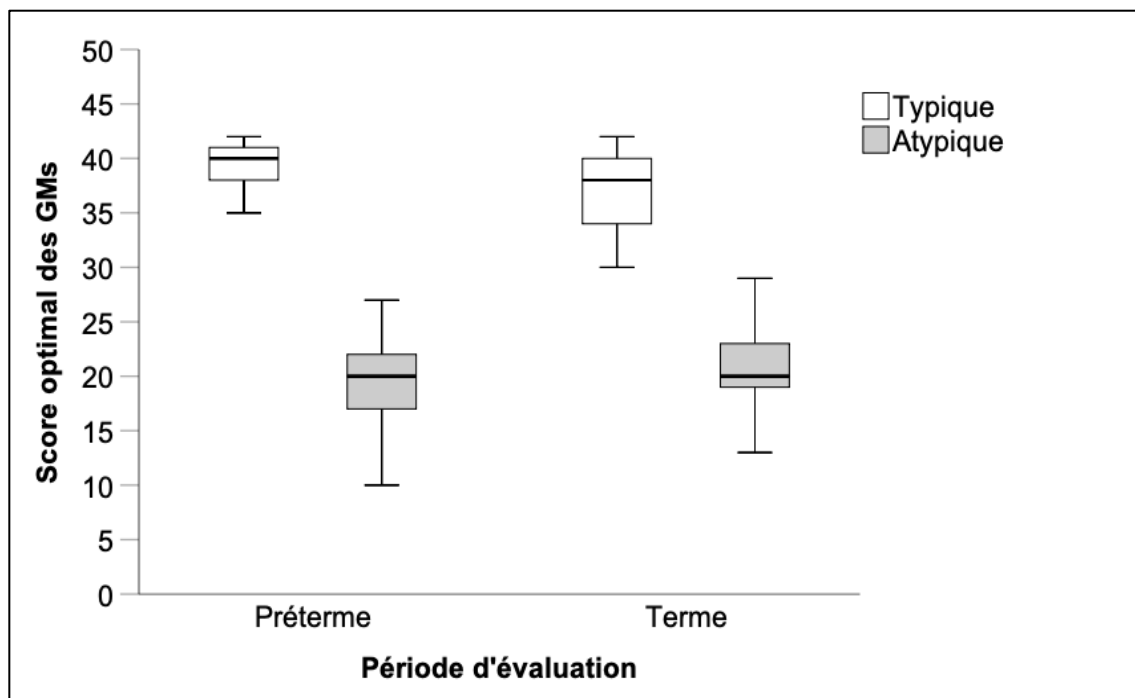
Figure 11. Score optimal des GMs à préterme et terme



Boîte aux moustaches représentant le score optimal des GMs à préterme et terme. La ligne à l'intérieur de la boîte indique la médiane des scores des GMs à préterme (Méd = 21) et terme (Méd = 24). Le haut et le bas de la boîte montrent l'intervalle où se situe le 50 % des scores à préterme (18–35) et terme (20–36). Les deux moustaches montrent le 25 % plus haut et le 25 % plus bas des scores à préterme (min = 10 ; max = 42) et terme (min = 13 ; max = 42). La valeur du score optimal des GMs à préterme et terme se situe entre 5 (indiquant la performance la moins bonne) et 42 (indiquant une performance optimale).

La figure 12 compare le score optimal des GMs typiques à celui des GMs atypiques à préterme et terme. À la période préterme, les scores des GMs typiques (Méd = 40) sont meilleurs que ceux des GMs atypiques (Méd = 21), (Kolmogorov-Smirnov) $Z = 3.638$, 95 % CI [18, 21], $p < .05$. De même, à terme, les scores des GMs typiques (Méd = 38) sont meilleurs que ceux des GMs atypiques (Méd = 20), $Z = 3.286$, 95 % CI [14, 19], $p < .05$, $r = .99$.

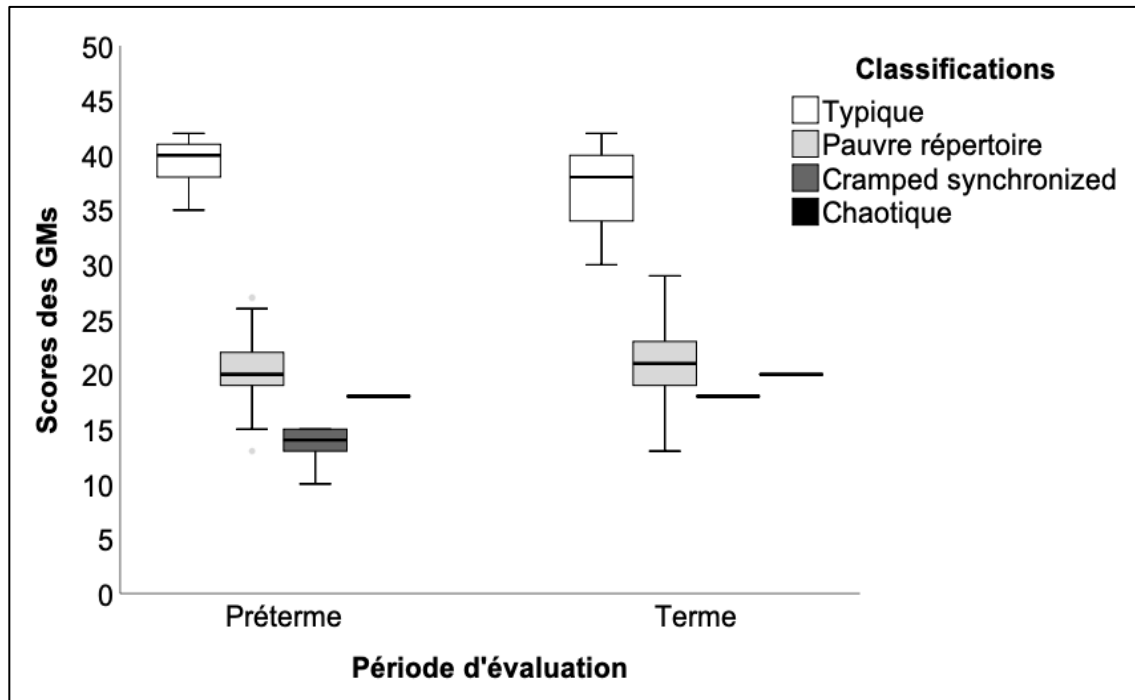
Figure 12. Score optimal des GMs typiques et GMs atypiques à préterme et terme



Boîte aux moustaches représentant le score optimal des GMs typiques et atypiques à préterme et terme. La ligne à l'intérieur de la boîte indique la médiane des scores des GMs typiques/atypiques à préterme (Méd = 40/20) et terme (Méd = 38/20). Le haut et le bas de la boîte montrent l'intervalle où se situe le 50 % des scores des GMs typiques/atypiques à préterme (37–41/17–22) et terme (33–40/19–23). Les deux moustaches montrent le 25 % plus haut et le 25 % plus bas des scores des GMs typiques/atypiques à préterme (min = 35/10 ; max = 42/27) et terme (min = 30/13 ; max = 42/29). La valeur du score optimal des GMs à préterme et terme se situe entre 5 (indiquant la performance la moins bonne) et 42 (indiquant une performance optimale).

La figure 13 compare le score optimal aux classifications des GMs dans la période préterme et terme. À préterme, les scores diffèrent significativement en fonction de la classification des GMs. En effet, la classification cramped synchronized implique les scores les plus bas (Méd = 14) par rapport aux autres classifications, (Jonckheere) $J = 21.500$, $p < .05$, $r = .92$. À terme, les scores diffèrent par rapport à la classification des GMs. Ainsi, la classification pauvre répertoire obtint les scores les plus bas (Méd = 21) comparés à ceux des autres classifications, $J = 22.020$, $p < .05$, $r = .81$.

Figure 13. Score optimal des classifications des GMs à préterme et terme



Boîte aux moustaches représentant le score optimal par rapport aux classifications des GMs à préterme et terme. La ligne à l'intérieur de la boîte indique la médiane des scores des GMs typiques/pauvre répertoire/cramped synchronized/chaotique à préterme (Méd = 40/20/14/13) et terme (Méd = 38/21/18/13). Le haut et le bas de la boîte montre l'intervalle où se situe le 50 % des scores des GMs typiques/pauvre répertoire/cramped synchronized/chaotique à préterme (37–41/19–22/12–15/13–13) et terme (33–40/19–23/18–18/13–13). Les deux moustaches montrent le 25 % plus haut et le 5 % plus bas des scores des GMs typiques/pauvre répertoire/cramped synchronized/chaotique à préterme (min = 35/13/10/13; max = 42/27/15/13) et terme (min = 30/13/18/13; max = 42/29/18/13). La valeur du score optimal des GMs à préterme et terme se situe entre 5 (indiquant la performance la moins bonne) et 42 (indiquant une performance optimale).

Le tableau 9 présente le pourcentage des GMs typiques et les GMs atypiques qui remplissent les 20 critères optimaux à préterme. D'une part, plus de 65 % des GMs typiques satisfont chaque critère optimal, sauf le critère « implication du cou » qui est faiblement rempli par les GMs typiques (34 %). D'autre part, moins de 25 % des GMs atypiques satisfont chaque critère optimal, excepté deux critères qui sont satisfaits par plus de 65 % des GMs atypiques : absence de cramped-synchronized aux bras (80 %) et jambes (70 %). En résumé, à préterme, chaque critère optimal est significativement plus satisfait par les GMs typiques que par les GMs atypiques, excepté quatre critères. Ainsi, les GMs typiques et les GMs atypiques ne diffèrent pas quant aux critères suivants : absence de tremblements aux bras, absence de cramped-synchronized aux bras, absence de tremblements aux jambes, absence de cramped-synchronized aux jambes.

Tableau 9. Pourcentage de GMs remplissant les critères optimaux à préterme

n ^a	TY 18	Méd	AT 50	Méd	Z	r
Cou impliqué	66 %	2	20 %	1	1.698*	.47
Rotations du tronc	33 %	1	0 %	1	2.506*	.69
Membres supérieurs						
Amplitude variable	94 %	2	8 %	1	3.145*	.87
Vitesse variable	83 %	2	0 %	1	3.032*	.84
Occupation spatiale	77 %	2	2 %	1	2.757*	.76
Rotations proximales	88 %	2	2 %	1	3.161*	.87
Rotations distales	94 %	2	8 %	1	3.145*	.87
Fluctuations au début	100 %	2	2 %	1	3.565*	.98
Fluctuations à la fin	100 %	2	4 %	1	3.493*	.96
Absence de tremblement	88 %	2	60 %	2	1.051	.29
Absence de cramped	100 %	2	88 %	2	.437	.19
Membres inférieurs						
Amplitude variable	94 %	2	0 %	1	3.436*	.95
Vitesse variable	88 %	2	0 %	1	3.234*	.89
Occupation spatiale	77 %	2	2 %	1	2.757*	.76
Rotations proximales	72 %	2	0 %	0	2.838*	.78
Rotations distales	94 %	2	0 %	0	3.436*	.95
Fluctuations au début	100 %	2	2 %	1	3.565*	.98
Fluctuations à la fin	100 %	2	4 %	1	3.493*	.96

Absence de tremblement	88 %	2	62 %	2	.978	.26
Absence de cramped	100 %	2	70 %	2	1.091	.31
Score optimal (n = 68)	26 %	40	73 %	21	3.638*	.99

^aUn enfant ne présente pas d'évaluation détaillée en raison d'une classification hypokinétique des GMs. TY, typique ; AT, atypique ; Méd, médiane ; Z, le test de Kolmogorov-Smirnov évalue la relation entre les données quantitatives non appariées ; r, le coefficient de Pearson quantifie l'effet sur la population : 0.1 = petit effet, 0.3 = effet moyen, 0.5 = grand effet ; *p < .05.

Le tableau 10 présente le pourcentage de GMs typiques et GMs atypiques qui remplissent les 20 critères optimaux à terme. D'un côté, plus de 70 % des GMs typiques remplissent chaque critère optimal, excepté cinq critères qui sont remplis par moins de 50 % des GMs typiques : rotations du tronc (45 %) ; occupation spatiale des bras en mouvement (34 %) ; variabilité de la vitesse des jambes (34 %) ; occupation spatiale des jambes en mouvement (45 %) ; rotation des hanches (34 %). De l'autre côté, moins de 55 % des GMs atypiques remplissent les critères optimaux, excepté trois critères qui sont remplis par plus de 70 % des GMs atypiques : absence de cramped-synchronized aux bras (92 %) ; absence de tremblements aux jambes (74 %) ; absence de cramped-synchronized aux jambes (74 %). En somme, à terme, chaque critère optimal est significativement plus satisfait par les GMs typiques que par les GMs atypiques, excepté quatre critères. Ainsi, les GMs typiques et les GMs atypiques ne diffèrent pas quant aux critères suivants : absence de tremblements aux bras, absence de cramped-synchronized aux bras, absence de tremblements aux jambes, absence de cramped-synchronized aux jambes.

Des comparaisons développementales montrent que les GMs à préterme et les GMs à terme ne diffèrent qu'en ce qui concerne quatre critères optimaux. Premièrement, les GMs à terme présentent une implication plus importante du cou que les GMs à préterme (Wilcoxon signed-rank), $T = 207.000$, $p < .05$, $r = .49$. Deuxièmement, l'implication du tronc est plus importante dans les GMs à terme que dans les GMs à préterme, $T = 145.000$, $p < .05$, $r = .36$. Troisièmement, les GMs à terme présentent une variabilité de la vitesse des bras plus importante que les GMs à préterme, $T = 102.000$, $p < .05$, $r = .32$. Quatrièmement, les mouvements des jambes

finalisent avec plus de souplesse et fluctuation en ce qui concerne les GMs à terme par rapport aux GMs à préterme, $T = 104.000$, $p < .05$, $r = .33$.

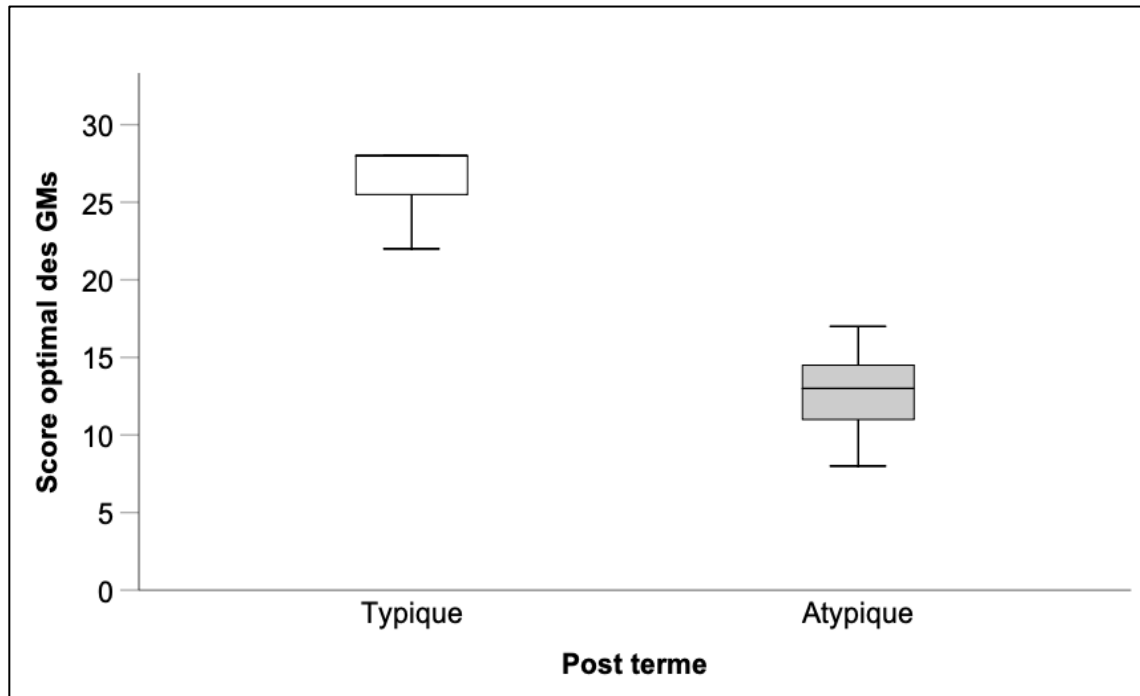
Tableau 10. Pourcentage de GMs remplissant les critères optimaux à terme

n	TY 18	Méd	AT 27	Méd	Z	r
Cou impliqué	94 %	2	40 %	1	1.765*	.53
Rotations du tronc	55 %	2	0 %	0	2.313*	.70
Membres supérieurs						
Amplitude variable	83 %	2	14 %	1	2.525*	.76
Vitesse variable	72 %	2	14 %	1	1.887*	.57
Occupation spatiale	66 %	2	22 %	1	1.461*	.44
Rotations proximales	94 %	2	0 %	1	3.104*	.94
Rotations distales	83 %	2	3 %	1	2.617*	.79
Fluctuations au début	83 %	2	3 %	1	2.617*	.79
Fluctuations à la fin	83 %	2	7 %	1	2.495*	.75
Absence de tremblement	77 %	2	51 %	2	.852	.25
Absence de cramped	100 %	2	92 %	2	.243	.07
Membres inférieurs						
Amplitude variable	72 %	2	14 %	1	2.008*	.61
Vitesse variable	66 %	2	7 %	1	1.947*	.59
Occupation spatiale	55 %	2	7 %	1	1.582*	.48
Rotations proximales	66 %	2	3 %	0	2.434*	.74
Rotations distales	72 %	2	3 %	0	2.252*	.68
Fluctuations au début	77 %	2	11 %	1	2.191*	.66
Fluctuations à la fin	83 %	2	11 %	1	2.373*	.72
Absence de tremblement	72 %	2	74 %	2	.304	.00
Absence de cramped	94 %	2	74 %	2	.669	.26
Score optimal (n = 45)	40 %	38	60 %	20	3.286*	.99

TY, typique ; AT, atypique ; Méd, médiane ; Z, le test de Kolmogorov-Smirnov évalue la relation entre les données quantitatives non appariées ; r, le coefficient de Pearson quantifie l'effet sur la population : 0.1 = petit effet, 0.3 = effet moyen, 0.5 = grand effet ; * $p < .05$.

La figure 14 compare le score optimal des GMs typiques à celui des GMs atypiques dans la période post terme. Ainsi, à post terme, les scores des GMs typiques (Méd = 28) sont plus élevés que les scores des GMs atypiques (Méd = 13), (Kolmogorov-Smirnov) $Z = 2.739$, 95 % CI [12, 15], $p < .05$, $r = .75$.

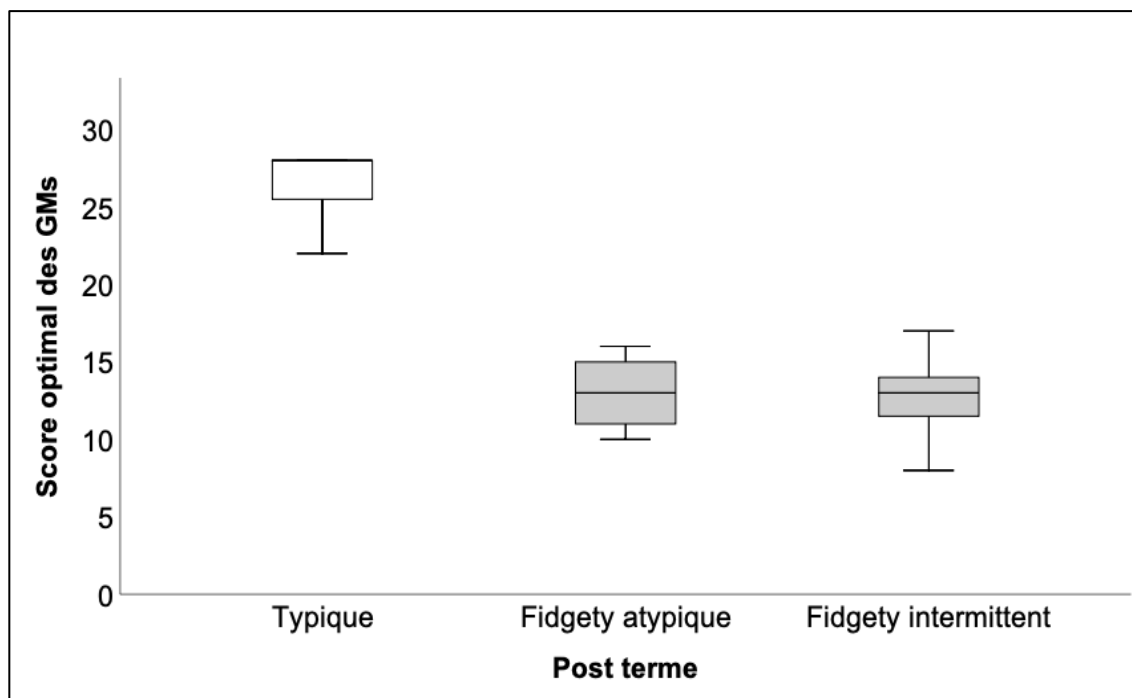
Figure 14. Score optimal des GMs à post terme



Boîte aux moustaches représentant le score optimal des GMs typiques et atypiques à post terme. La ligne à l'intérieur de la boîte indique la médiane des scores des GMs typiques/atypiques (Méd = 28/13). Le haut et le bas de la boîte montre l'intervalle où se situe le 50 % des scores des GMs typiques/atypiques (25–28/10–14). Les deux moustaches montrent le 25 % plus haut et le 25 % plus bas des scores des GMs typiques/atypiques (min = 22/8 ; max = 28/17). La valeur du score optimal des GMs à post terme se situe entre 5 (indiquant la performance la moins bonne) et 28 (indiquant une performance optimale).

La figure 15 compare le score optimal aux classifications des GMs à post terme. À post terme, la classification des GMs détermine les scores. En effet, la classification fidgety intermittent implique les scores les plus bas (Méd = 13) par rapport aux autres classifications des GMs à post terme, (Jonckheere) $J = .000$, $p < .05$, $r = .75$.

Figure 15. Score optimal des classifications des GMs à post terme



Boîte aux moustaches représentant le score optimal par rapport aux classifications des GMs à la période post terme. La ligne à l'intérieur de la boîte aux moustaches indique la médiane des scores des GMs typiques/fidgety atypique/fidgety intermittent (Méd = 28/13/13). Le haut et le bas de la boîte montre l'intervalle où se situe le 50 % des scores des GMs typiques/fidgety atypique/fidgety intermittent (25–28/10–15/10–14). Les deux moustaches montrent le 25 % plus haut et le 25 % plus bas des scores des GMs typiques/fidgety atypique/fidgety intermittent (min = 22/10/8 ; max = 28/16/17). La valeur du score optimal des GMs à post terme se situe entre 5 (indiquant la performance la moins bonne) et 28 (indiquant une performance optimale).

Le tableau 11 présente le pourcentage des GMs typiques et GMs atypiques qui remplissent les 5 critères optimaux à post terme. D'une part, plus de 80 % des GMs typiques satisfont chaque critère optimal, sauf le critère « taille du répertoire d'autres mouvements accompagnant les GMs » étant faiblement rempli par les GMs typiques (35 %). D'autre part, moins de 60 % des GMs atypiques satisfont chaque critère optimal, excepté le critère « qualité du répertoire d'autres mouvements accompagnant les GMs » étant rempli par plus de 70 % des GMs atypiques. En résumé, bien que chaque critère optimal soit plus atteint par les GMs typiques que par les GMs atypiques, la seule différence statistiquement significative concerne au

critère suivant : caractère de la motricité. En d'autres termes, à post terme, le caractère souple et fluide de la motricité est rencontré significativement plus souvent pour les GMs typiques que pour les GMs atypiques.

Tableau 11. Pourcentage de GMs remplissant les critères optimaux à post terme

n	TY 20	Méd	AT 12	Méd	Z	r
Répertoire d'autres mouvements	65 %	4	16 %	2	1.324	.48
Qualité d'autres mouvements	100 %	4	75 %	4	.685	.25
Posture typique	85 %	4	58 %	4	.730	.26
Motricité souple et fluide	90 %	4	16 %	2	2.008*	.73
Score optimal (n = 32)	62 %	28	37 %	13	2.739*	.75

TY, typique ; AT, atypique ; Méd, médiane ; Z, le test de Kolmogorov-Smirnov évalue la relation entre les données quantitatives non appariées ; r, le coefficient de Pearson quantifie l'effet sur la population : 0.1 = petit effet, 0.3 = effet moyen, 0.5 = grand effet ; *p < .05.

6.3. Discussions

Afin de constituer une base de données de GMs, en tant que mesure pour développer un outil neuromoteur, cette thèse a longitudinalement évalué les GMs des enfants nés prématurément. Plus particulièrement, cette étude décrit la mesure par rapport à : la détection des changements développementaux relatifs à la neuromotricité ; la description de la trajectoire clinique de la neuromotricité ; la distinction entre les enfants ayant des anomalies neuromotrices et ceux n'en ayant pas. Trois résultats principaux ont été obtenus.

Premièrement, le diagnostic (typique/atypique) et le score optimal des GMs varient quant à la détection de changements développementaux. Tout d'abord, le diagnostic des GMs n'a pas permis d'identifier de changements entre les trois périodes d'évaluation. Cela pourrait s'expliquer par le fait que le diagnostic des GMs,

au lieu de détecter des changements liés au développement, se rapporte à des aspects cliniques pour détecter les anomalies neuromotrices. Ensuite, le score optimal des GMs à préterme ou terme ne prédit pas le score optimal des GMs à post terme. Ce résultat peut s'expliquer par la grande fréquence de la classification pauvre répertoire à préterme et terme, dont la capacité à prédire le développement n'est pas prouvée (Nakajima, Einspieler, Marschik, Bos, & Prechtl, 2006) ; Prechtl et al., 1997). Après, le score optimal des GMs à préterme prédit celui des GMs à terme, car un bon score à préterme détermine un bon score à terme. De même, le score optimal des GMs s'améliore à terme. À l'opposé, une étude a montré que le score optimal des GMs ne s'améliore pas, voire qu'il diminue à terme (Einspieler et al., 2015). Cependant, l'étude mentionnée considère une tranche d'âge (entre 37 SAGc et 45 SAGc) plus large que la nôtre (entre 37 SAGc et 41 SAGc). Ainsi, la diminution du score optimal que l'étude mentionnée a observé à terme reflèterait, au lieu d'une différence développementale, la transition d'une période à l'autre : celle des GMs writhing disparaissant et celle des GMs fidgety commençant. Enfin, comparé à la période préterme, nous avons observé des améliorations à la période terme concernant les critères optimaux des GMs (composants de l'évaluation détaillée des GMs. Voir les tableaux 9 et 10). Ces améliorations se rapportent aux critères suivants : une implication plus importante du cou et tronc ; une plus grande variabilité de la vitesse des bras ; une fluidité plus importante dans les jambes en mouvement. À contrario, des études concluent que le tronc présente une implication plus notable à la période préterme qu'à terme (Hadders-Algra & Prechtl, 1992 ; Hadders-Algra et al., 1997). Ces études se réfèrent aux changements développementaux qui sont survenus sur un échantillon de GMs typiques, alors que notre étude reflète des changements sur un échantillon hétérogène. De plus, notre étude a observé une implication importante du tronc et du cou en termes qualitatifs, à savoir des rotations plus complexes à la période terme comparé à celles de la période préterme. En effet, ce même résultat a été observé dans une étude longitudinale (Zorzenon, Takaara, & Linhares, 2019). En outre, des études confirment nos résultats concernant l'augmentation de la variabilité de la vitesse des GMs avec l'augmentation de l'âge gestationnel corrigé (Cioni & Prechtl, 1990).

Cependant, ces résultats pourraient aussi représenter une amélioration clinique à terme, car une vitesse monotone est caractéristique des enfants à risque (Bos, van Asperen, de Leeuw, & Prechtl, 1997; Touwen, 1990). L'amélioration de la souplesse avec l'âge gestationnel corrigé n'est pas confirmée dans d'autres études.

Deuxièmement, le diagnostic (typique/atypique) et le score optimal des GMs diffèrent quant à la description de la trajectoire clinique neuromotrice. D'une part, cette étude montre que le diagnostic des GMs s'améliore progressivement, car la proportion des GMs atypiques diminue à travers les trois périodes d'évaluation. Ce résultat a été confirmé par des études qui observent une grande prévalence des GMs atypiques à préterme et terme, dont la normalisation se constate par la prédominance des GMs typiques à post terme (Albers, 1994; Bruggink, Koenraad, Braeckel, & Bos, 2010; Cioni et al., 1997; de Vries & Bos, 2011; de Vries & Bos, 2010; de Vries, Erwich, & Bos, 2008; Olsen et al., 2015; Zahed-Cheikh et al., 2011). D'autre part, bien que le score optimal des GMs atypiques ne s'améliore pas entre la période préterme et terme, des améliorations cliniques s'observent concernant les critères optimaux des GMs (composants de l'évaluation détaillée des GMs. Voir les tableaux 9 et 10). Ainsi, le pourcentage des GMs atypiques qui atteignent les critères optimaux augmente à terme (6 %) par rapport à préterme. Cependant, les tremblements et les composants cramped-synchronized ne permettent pas d'identifier des différences cliniques à préterme et terme en raison de deux constats. Premièrement, bien que de courte durée, les tremblements peuvent être observés également dans les GMs typiques : ce ne sont donc pas une exclusivité des GMs atypiques. Deuxièmement, concernant surtout les bras, les composants cramped synchronized sont très peu observés pour les GMs atypiques et nullement observés pour les GMs typiques. Même si les composants cramped synchronized s'observent plus fréquemment sur les jambes, ils ne le sont pas encore assez pour départager les GMs typiques des GMs atypiques.

Troisièmement, le score optimal des GMs diffère entre les enfants ayant des anomalies motrices (GMs atypiques) et ceux n'en ayant pas (GMs typiques). En effet, sur les trois périodes d'évaluation, les scores des GMs typiques sont meilleurs que ceux des GMs atypiques. Ce résultat confirme donc le diagnostic des GMs et la

composition des deux groupes : celui des GMs typiques et celui des GMs atypiques. Dans ce sens, des études démontrent que le score optimal des GMs aide à identifier les GMs atypiques et les GMs typiques (Einspieler et al., 2015; Ustad, Evensen, Bertocelli, Frassoldati, & Ferrari, 2017).

Encadré 6. Résumé du chapitre VI

Un outil automatisé requiert d'être programmé avec un modèle d'activité étant construit à partir d'une mesure neuromotrice. Cette première étude concerne la constitution d'une mesure neuromotrice qui, sous forme de base de données des GMs, fut obtenue auprès de 75 enfants nés prématurément. Les objectifs de cette étude visent la description des GMs en tant que mesure neuromotrice permettant de : détecter les changements neuromoteurs qui surviennent au cours du développement ; tracer la trajectoire clinique de la neuromotricité ; distinguer entre les enfants ayant des anomalies neuromotrices et ceux n'en ayant pas. La motricité spontanée de l'échantillon fut filmée et évaluée à trois périodes : préterme, terme et post terme. Au total, 149 enregistrements vidéo furent obtenus dont 785 séquences des GMs furent extraites : 347 à préterme, 257 au terme, 181 à post terme. Trois résultats principaux ont été obtenus. Premièrement, le diagnostic (typique/atypique) des GMs et le score optimal des GMs diffèrent quant à l'identification des changements développementaux. D'une part, le diagnostic n'a pas permis d'identifier de changements développementaux. De même, le score à préterme ne prédit pas le score à post terme. D'autre part, le score à préterme prédit celui à terme concernant les améliorations suivantes : une implication plus importante du cou et tronc ; une grande variabilité de la vitesse des bras ; une fluidité plus importante dans les jambes en mouvement. Deuxièmement, le diagnostic des GMs et le score des GMs diffèrent quant à la description de l'évolution clinique de la neuromotricité. D'un côté, le nombre des GMs atypiques diminue à travers les trois périodes d'évaluation, de l'autre le score des GMs atypiques ne s'améliore pas à terme. Pourtant, des améliorations cliniques s'observent concernant les critères optimaux des GMs à terme. Troisièmement, le score des GMs diffère entre les enfants ayant des anomalies motrices et ceux n'en ayant pas, car le score des GMs atypiques est moins élevé que celui des GMs typiques. Bien que ces résultats soient prometteurs, car ils décrivent la base de données de GMs, leur adéquation doit être évaluée par des études de validation. Le chapitre suivant évalue l'adéquation des GMs, en tant que mesure neuromotrice, pour le nouveau-né prématuré qui est hospitalisé.

VII. Validation d'une base de données des mouvements généraux

Les études de validation des GMs présentent trois caractéristiques qui rendent difficile de juger de l'adéquation des GMs pour le nouveau-né prématuré qui est à l'hôpital. Premièrement, la plupart des études de validation considère les GMs quand l'enfant a quitté l'hôpital, entre le 3^e et 5^e mois de vie (Adde, Rygg, Lossius, Øberg, & Støen, 2007 ; Bosanquet et al., 2013 ; Constantinou et al., 2007 ; Darsaklis, Snider, Majnemer, & Mazer, 2011 ; Fjørtoft et al., 2009 ; Paro-Panjan, Sustersic, & Neubauer, 2005 ; Seme-Ciglencecki, 2007 ; Snider, Majnemer, Mazer, Campbell, & Bos, 2009 ; Spittle, Boyd, Inder, & Doyle, 2009). Deuxièmement, des études de validation ne se focalisent pas sur les GMs à la période préterme, car elles examinent plusieurs périodes développementales confondues (Ferrari et al., 2002 ; Garcia, Gherpelli, & Leone, 2004 ; Guzzetta et al., 2007 ; Rau, Lin, Ko, & Hsue, 2014 ; Yeh et al., 2016b). Troisièmement, la plupart des études de validation des GMs n'a pas été conçue en tant que telle (Annexe 3) (Nakajima et al., 2006 ; Olsen et al., 2016 ; Olsen et al., 2015 ; Roescher et al., 2014). En effet, l'organisation des services de néonatalogie et les particularités du nouveau-né prématuré rendent difficile la réalisation des études de validation des GMs sur cette population. De même, les mesures automatisées des GMs, à notre connaissance, ne présentent pas d'études de validation sur le nouveau-né prématuré. Cela s'explique en partie par le fait que les mesures automatisées, bien qu'elles soient destinées aux enfants nés prématurés, ne devraient être utilisées qu'après le terme.

Par conséquent, l'objectif de cette thèse concerne la validation d'une base de données sur les GMs qui, en tant que mesure neuromotrice, permettra le développement d'un outil automatisé destiné au nouveau-né prématuré qui est hospitalisé. Plus particulièrement, cette deuxième étude évalue quatre propriétés

psychométriques des GMs sur cette population : la validité du construit, la validité de critère, la fiabilité intracodeur et la fiabilité intercodeur. La validité du construit évalue le degré auquel une mesure reflète la construction théorique d'intérêt, dans notre cas la fonction neuromotrice du nouveau-né prématuré (Fermanian, 2005). Le paramètre de la validité du construit est la relation entre les scores de la mesure neuromotrice en question et les adversités cliniques qui, survenues dans la période prénatale, périnatale ou néonatale, représentent un risque neurologique (Heineman & Hadders-Algra, 2008 a). La validité des critères évalue le degré auquel une mesure reflète les scores du test considéré comme le gold standard pour la neuromotricité chez le nouveau-né prématuré. Dans l'absence d'un tel test, la validité des critères s'élargit pour désigner un diagnostic spécifique comme la mesure de référence (Roach, 2006). Dans ce cas, le diagnostic d'une lésion cérébrale étant détectée par échographie transfontanellaire fut utilisé comme paramètre de la validité de critère (Heineman et al., 2010). La fiabilité intracodeur est le degré d'accord entre des évaluations répétées qui sont réalisées par un même codeur sur le même patient, stable, à différentes occasions. La fiabilité intercodeur est le degré d'accord entre des évaluations répétées qui sont réalisées par plusieurs codeurs, en même temps, sur le même patient (Lohr et al., 1996 ; Mokkink et al., 2010 ; Terwee et al., 2007).

Cette étude émet quatre hypothèses quant aux GMs sur le nouveau-né prématuré : la validité du construit se vérifie par une relation significative entre les GMs et des variables cliniques qui représentent un risque ; les épreuves sur la validité des critères se réfèrent à une bonne capacité des GMs à détecter les lésions cérébrales ; les GMs présentent une bonne fiabilité intracodeur ; les GMs présentent une bonne fiabilité intercodeur.

7.1. Méthodologie

La méthodologie générale de cette étude a été décrite dans le chapitre VI.

7.1.1. Participants

Le processus de recrutement est décrit dans le chapitre VI (voir la figure 8). Quatre-vingt-deux enfants furent recrutés, dont 69 enfants constituent l'échantillon final. Vingt-cinq pourcent ($n = 21$) des enfants provenaient d'une grossesse multiple et 49 % ($n = 38$) étaient des garçons. Le tableau 12 détaille les caractéristiques de l'échantillon, dont la prématurité se distribue en : 5 % ($n = 4$) des enfants sont nés extrêmement prématurés, 43 % ($n = 26$) nés grands prématurés, 45 % ($n = 37$) nés moyennement ou tardivement prématurés. Au moins une complication néonatale a été observée chez 41 % ($n = 34$) des enfants, dont 35 % ($n = 29$) présentaient une échographie cérébrale atypique. L'âge au moment de l'évaluation des GMs se situe entre 31.2 AGc et 36.8 AGc.

7.1.2. Procédure

Le chapitre VI décrit la procédure de recueil et évaluation des GMs. AV (codeur 1) a codé les GMs (codage 1) après d'avoir relevé les données cliniques dans les dossiers d'hospitalisation des enfants. Un an après, le codeur 1 a recodé les GMs (codage 2). Cet intervalle de temps a été considéré comme suffisamment long pour minimiser le biais de mémoire du codeur 1 (Jeng et al., 2000). Pour cette deuxième étude, JP (codeur 2) a indépendamment réalisé l'analyse globale des GMs sur 100 % ($n = 69$) de l'échantillon et l'analyse détaillée des GMs sur 51 % ($n = 35$) de l'échantillon. JP ne connaissait pas les caractéristiques cliniques des enfants lors du codage des GMs.

Tableau 12. Caractéristiques de l'échantillon (n = 69)

Données cliniques

Âge gestationnel (SA), moyenne (SD)	31.6 (2.3)
Poids de naissance (g), moyenne (SD)	1534.5 (413.5)
APGAR à 1 min, moyenne (SD)	7 (2)
APGAR à 3 min, moyenne (SD)	8 (2)
APGAR à 5 min, moyenne (SD)	8 (1.4)
APGAR à 10 min, moyenne (SD)	9 (0.9)
Poids extrêmement faible à la naissance, n (%)	7 (8)
Petit pour l'âge gestationnel, n (%)	16 (19)
Périmètre crânien à la naissance (cm), moyenne (SD)	28.5 (2.3)
Hémorragie intraventriculaire, n (%)	13 (15)
I, n (%)	5 (6)
II, n (%)	6 (7)
III, n (%)	2 (2)
Leucomalacie périventriculaire, n (%)	24 (29)
I, n (%)	18 (22)
II, n (%)	4 (5)
III, n (%)	2 (2)
Dysplasie bronchopulmonaire, n (%)	12 (14)
Légère, n (%)	7 (8)
Modérée, n (%)	5 (6)
Entéocolite ulcéro-nécrosante, n (%)	1 (1)
Infection, n (%)	13 (16)
Bradycardie, n (%)	20 (24)
Apnée, n (%)	21 (25)
Détresse respiratoire	49 (60)
Légère, n (%)	23 (28)
Modérée, n (%)	26 (32)
Ventilation artificielle, n (%)	54 (66)
Ventilation invasive, n (%)	44 (54)
Ventilation invasive (jours), moyenne (SD)	2.4 (5)
Ventilation non invasive, n (%)	50 (66)
Ventilation non invasive (jours), moyenne (SD)	11.6 (18)

Mouvements généraux

Atypiques, n (%)	51 (73)
Score optimal ^a , moyenne (SD)	24 (9)

SD, écart-type ; APGAR, le score d'APGAR reflète l'état du nouveau-né à la minute 1, 3, 5, 7 et 10 se suivant à la naissance. Le score d'APGAR évalue cinq critères : la fréquence cardiaque, l'effort respiratoire, le tonus musculaire, l'irritabilité et la couleur de la peau. La somme de ces cinq critères produit le score dont la valeur de 7 ou plus indique que l'état du bébé est bon (Casey, McIntire, & Leveno, 2001) ; Petit pour l'âge gestationnel : défini comme le poids à la naissance par rapport à l'âge gestationnel inférieur au 10^e percentile (Chabra, 2018) ; Dysplasie bronchopulmonaire : définie comme le besoin d'oxygène supplémentaire pendant au moins 28 jours à 36 SAGc chez le grand prématuré, et à l'âge chronologique de 56 jours chez l'enfant né

modérément prématuré (Hadchouel & Delacourt, 2013 ; Sherlock & Abman, 2018) ; ^aLa valeur du score optimal des GMs se situe entre 5 (indiquant la performance la moins bonne) et 42 (indiquant une performance optimale).

Le résultat de la dernière échographie cérébrale a été pris en compte et catégorisé comme typique ou atypique. Un résultat atypique était classé selon la gravité du diagnostic de l'HIV (hémorragie intraventriculaire) et la LPV (leucomalacie périventriculaire). L'HIV se classe en quatre stades qui augmentent de I à IV en fonction de la gravité de la lésion : le stade I se caractérise par un saignement dans la zone germinative ; le stade II implique une HIV sans dilatation ventriculaire ; le stade III comporte une HIV avec augmentation aiguë du ventricule ; le stade IV comporte une congestion et une lésion du tissu périventriculaire (Ment et al., 2012). De même, la gravité de la LPV se classe en quatre niveaux : le niveau I se caractérise par une hyperéchogénicité périventriculaire transitoire qui dure plus de 7 jours ; le niveau II présente une hyperéchogénicité périventriculaire transitoire pendant plus de 7 jours et/ou des petits kystes ; le niveau III implique une hyperéchogénicité périventriculaire avec des lésions kystiques étendues ; le niveau IV comprend une hyperéchogénicité dans la substance blanche, avec des lésions kystiques étendues (de Vries, Eken, & Dubowitz, 1992; McCrea, Ment, & McCrea, 2010). L'échographie cérébrale a été régulièrement réalisée par des pédiatres radiologues qui ignoraient le déroulement de cette étude et selon le mode d'organisation du service concerné.

7.1.3. Traitement et analyse des données

Le codage 2 fait l'objet de validation dans cette étude, car : le codeur 1 avait plus d'expérience au moment du codage 2 ; il présente une bonne fiabilité ; il est décrit dans l'étude rapportée au chapitre VI. Le chapitre VI définit les trois mesures des GMs qui seront validées dans cette étude. Premièrement, le diagnostic

(typique/atypique) des GMs obtenu lors de l'évaluation globale. Deuxièmement, le score optimal des GMs obtenu lors de l'évaluation détaillée. Le score optimal se situe entre 5 et 42 indiquant la performance la moins bonne et une performance optimale, respectivement. Troisièmement, les sous-scores des 20 critères optimaux qui sont considérés dans l'évaluation détaillée des GMs (voir le tableau 9).

La validité du construit a été évaluée par l'analyse de régression. L'analyse de régression logistique a été utilisée pour mesurer la relation entre le diagnostic (typique/atypique) des GMs et les variables cliniques. L'analyse de régression linéaire a été utilisée pour évaluer la relation entre le score optimal des GMs et les variables cliniques. Dans ces deux méthodes, les variables cliniques furent d'abord évaluées par un modèle univarié, afin d'identifier des associations significatives, pour ensuite les combiner dans un modèle multivarié qui identifie des relations indépendantes. Le modèle multivarié fut construit par une méthode pas à pas (backward stepwise). La variable EUCN (entérocolite ulcéro-nécrosante) ne fut pas considérée à cause de données insuffisantes pour l'analyse de régression. La validité des critères fut évaluée par des indices épidémiologiques et l'analyse de régression. La sensibilité, la spécificité, la VPP (valeur prédictive positive) et la VPN (valeur prédictive négative) furent utilisés pour mesurer la précision diagnostique des GMs par rapport à l'HIV (hémorragie intraventriculaire) et la LPV (leucomalacie périventriculaire). Le tableau 3, dans le chapitre IV, décrit ces indices épidémiologiques. L'analyse de régression linéaire a été utilisée pour mesurer la relation du score optimal des GMs avec l'HIV et la LPV.

La fiabilité a été évaluée par des mesures d'accord. Le k (kappa) a été utilisé pour évaluer l'accord intracodeur et intercodeur sur le diagnostic (typique/atypique) des GMs. Selon les standards de (McHugh, 2012) : une valeur de k entre 0.21 et 0.39 montre une fiabilité minimale ; une valeur comprise entre 0.40 et 0.59 indique une fiabilité faible ; une valeur entre 0.60 et 0.79 montre une fiabilité modérée ; une valeur entre 0.80 et 0.90 représente une fiabilité forte ; une valeur au-dessus de 0.90 représente une fiabilité presque parfaite. L'ICC (coefficient de corrélation intraclasse) a été utilisé pour évaluer l'accord intracodeur et intercodeur sur le score optimal des GMs. Selon les standards de Cicchetti, (1994) : une valeur de l'ICC

comprise entre 0.40 et 0.59 montre une fiabilité modérée ; une valeur entre 0.60 et 0.74 indique une bonne fiabilité ; une valeur comprise entre 0.75 et 1.0 révèle une fiabilité excellente. L'accord intracodeur pour le diagnostic (typique/atypique) des GMs et le score optimal des GMs furent évalués sur 100 % (n = 69) de l'échantillon. L'accord intercodeur pour le diagnostic des GMs fut évalué sur 100 % (n = 69) de l'échantillon et l'accord intercodeur pour le score optimal des GMs fut évalué sur 51 % (n = 35) de l'échantillon.

La valeur de $p < .05$ était considérée statistiquement significative et accompagnée d'un IC BCa (intervalle de confiance sur la méthode bootstrap) fixé à 95 %. L'ampleur de l'effet est rapportée par l'OR (odds ratio). Les analyses statistiques furent réalisées avec IBM SPSS Statistics. Les indices épidémiologiques furent calculés sur le site web : https://www.medcalc.org/calc/diagnostic_test.php.

7.2. Résultats

7.2.1. Validité

Le tableau 13 montre la relation entre le diagnostic (typique/atypique) des GMs et des variables cliniques qui représentent un risque. Dans la régression univariée, le diagnostic des GMs présente une relation significative avec l'âge gestationnel, le score d'APGAR à 5 min, au moins une complication néonatale, l'échographie cérébrale et le nombre de jours d'hospitalisation. D'une part, l'OR (odds ratio) pour un diagnostic de GMs atypiques diminue avec l'augmentation de l'âge gestationnel et l'augmentation du score d'APGAR à 5 min. D'autre part, l'OR pour un diagnostic de GMs atypiques augmente avec une échographie cérébrale atypique, au moins une complication néonatale et le nombre de jours d'hospitalisation. Dans le modèle

multivarié, le diagnostic des GMs n'est indépendamment associé qu'au diagnostic de l'échographie cérébrale.

Tableau 13. Régression logistique pour le diagnostic des GMs (n = 69^a)

	Coefficient de régression	Erreur standard	OR	95 % IC de l'OR	
				Inférieur	Supérieur
Univariée					
Âge gestationnel, SA	-0.38*	0.14	0.68	0.48	0.94
APGAR à 5 min	-0.77*	0.38	0.46	0.21	0.98
Complication néonatale ^b	1.42*	0.63	4.13	1.18	14.46
Échographie cérébrale	1.54*	0.70	4.69	1.19	18.52
Jours d'hospitalisation	0.03*	0.01	1.03	1.00	1.07
Multivariée					
Constante	0.69	0.35			
Échographie cérébrale	2.25*	0.17	9.50	1.13	79.69

La régression logistique évalue la relation entre une variable dichotomique (ex. diagnostic typique/atypique des GMs) et des prédicteurs (ex. données cliniques); ^aTrois enfants n'avaient pas de données concernant l'échographie cérébrale; OR, l'odds ratio compare la cote des GMs atypiques sur un groupe A (ex. enfants ayant une échographie cérébrale atypique) avec la cote des GMs atypiques sur un groupe B (ex. enfants ayant une échographie cérébrale étant typique). Plus l'OR s'écarte de 1.0, plus la différence entre les groupes est importante; IC, l'intervalle de confiance est une plage qui contient la valeur réelle sur la population; *p < .05; ^bDysplasie broncho-pulmonaire, entérocolite ulcéro-nécrosante, infection néonatale et/ou lésion cérébrale.

Le tableau 14 détaille la relation entre le score optimal des GMs et des variables cliniques qui représentent un risque. Dans la régression univariée, le score optimal des GMs augmente avec l'augmentation de l'âge gestationnel et l'augmentation du score d'APGAR à 3 min et à 5 min. De même, le score optimal des GMs diminue avec : au moins une complication néonatale ; une échographie cérébrale atypique ; l'HIV (hémorragie intraventriculaire) ; la gravité de l'HIV ; la LPV (leucomalacie périventriculaire) ; la gravité de la LPV ; la DPB (dysplasie bronchopulmonaire) ; la gravité de la DBP ; la ventilation artificielle invasive et non invasive ; les jours

d'hospitalisation. Dans le modèle multivarié, le score optimal des GMs n'est indépendamment associé qu'au score d'APGAR à 3 min et à l'HIV (tous les niveaux de gravité confondus), $F(2,51) = 6.877$, $p < .05$. La valeur de R^2 indique que 21 % de la variation du score optimal des GMs s'explique par le score d'APGAR à 3 min et par l'HIV.

Tableau 14. Régression linéaire pour le score optimal des GMs (n =69^a)

	Coefficient de régression	Erreur standard	95 % IC	
			Inférieur	Supérieur
Univariée				
Âge gestationnel, SA	1.46*	0.47	0.51	2.41
APGAR à 3 min	1.24*	0.57	0.09	2.39
APGAR à 5 min	2.23*	0.79	0.63	3.83
Complication néonatale ^b	-5.85*	2.19	-10.23	-1.47
Échographie cérébrale	-6.25*	2.19	-10.63	-1.87
Hémorragie intraventriculaire (HIV)	-6.57*	2.76	-12.06	-1.05
Gravité de l'HIV	-3.46*	1.42	-6.31	-0.61
Leucomalacie périventriculaire (LPV)	-5.15*	2.32	-9.79	-0.50
Gravité de la LPV	-3.94*	1.46	-6.86	-1.02
Dysplasie bronchopulmonaire (DBP)	-7.87*	2.99	-13.85	-1.89
Gravité de la DBP	-5.70*	2.03	-9.73	-1.64
Ventilation artificielle, jours	-0.13*	0.05	-0.23	-0.02
Ventilation invasive, n	-5.10*	2.25	-9.61	-0.59
Ventilation invasive, jours	-0.48*	0.21	-0.90	-0.06
Ventilation non invasive, jours	-0.13*	0.06	-0.25	-0.00
Jours d'hospitalisation	-0.15*	0.05	-0.25	-0.04
Multivariée				
Constante	14.34*	2.77		
APGAR à 3 min	1.34*	0.57	0.19	2.50
Hémorragie intraventriculaire	-6.16*	2.77	11.74	-0.59

La régression linéaire évalue la relation entre une variable quantitative (ex. score optimal des GMs) et des prédictors (ex. données cliniques) ; ^aTrois enfants n'avaient pas de données concernant l'échographie cérébrale ; IC, l'intervalle de confiance est une plage qui contient la valeur réelle sur la population ; * $p < .05$; ^bDysplasie broncho-pulmonaire, entérocolite ulcéro-nécrosante, infection néonatale et/ou lésion cérébrale.

Le tableau 15 montre la relation entre le diagnostic (typique/atypique) des GMs et le diagnostic pour l'HIV (hémorragie intraventriculaire). Les GMs atypiques présentent une sensibilité de 92 % [95 % IC = 64 %, 100 %] pour détecter les enfants ayant un diagnostic positif pour l'HIV (tous les niveaux de gravité confondus). Les GMs typiques présentent une spécificité de 28 % [95 % IC = 17 %, 42 %] pour détecter les enfants qui ne présentent pas d'HIV. La VPP (valeur prédictive positive) est 24 % [95 % IC = 20 %, 28 %] et la VPN (valeur prédictive négative) est 94 % [95 % IC = 68 %, 94 %].

Tableau 15. Diagnostic des GMs et l'hémorragie intraventriculaire

GMs	HIV			Total
	Non	I	II – III	
Typique	15	1	0	16
Atypique	38	4	8	50
Total	53	5	8	66

Le tableau 16 détaille la relation entre le diagnostic (typique/atypique) des GMs et le diagnostic pour la LPV (leucomalacie périventriculaire). Les GMs atypiques montrent une sensibilité de 87 % [95 % IC = 68 %, 97 %] pour détecter les enfants ayant un diagnostic positif pour la LPV (tous les niveaux de gravité confondus). Les GMs typiques montrent une spécificité de 31 % [95 % IC = 18 %, 47 %] pour identifier les enfants qui ne présentent pas de LPV. La VPP (valeur prédictive positive) est 42 % [95 % IC = 36 %, 48 %] et la VPN (valeur prédictive négative) est 81 % [95 % IC = 58 %, 93 %].

Tableau 16. Diagnostic des GMs et la leucomalacie périventriculaire

GMs	LPV			Total
	Non	I	II – III	
Typique	13	3	0	16
Atypique	29	15	6	50
Total	42	18	6	66

Le tableau 17 montre la relation entre le score optimal des GMs et le diagnostic pour l'HIV ou la LPV. Le score optimal des GMs et le diagnostic de l'HIV ou de la LPV présentent une relation significative, $F(2,62) = 4.299$, $p < .05$. En effet, le score optimal des GMs diminue avec un diagnostic positif pour l'HIV ou la LPV (tous les stades ou degrés de gravité confondus). Ainsi, la valeur de R^2 indique que 34 % de la variation du score optimal des GMs s'explique par une de ces deux lésions cérébrales. Cependant, la gravité de la lésion cérébrale (HIV ou LPV) n'est pas associée au score optimal des GMs, car elle n'explique pas la variation de celui-ci.

Tableau 17. Régression linéaire pour le score des GMs et l'HIV ou la LPV ($n = 69^a$)

	Coefficient de régression	Erreur standard	95 % IC de BCa	
			Inférieur	Supérieur
HIV	-5.28*	2.04	-8.89	-1.43
LPV	-3.94	2.20	-8.02	-0.09

La régression linéaire évalue la relation entre une variable quantitative (ex. score optimal des GMs) et des prédicteurs (ex. HIV et LPV) ; ^aTrois enfants n'avaient pas de données concernant l'échographie cérébrale ; IC de BCa, l'intervalle de confiance est une plage qui contient la valeur réelle sur la population. La méthode bootstrap (BCa) permet de calculer l'IC en corrigeant le biais ; * $p < .05$.

7.2.2. Fiabilité

Le tableau 18 montre le pourcentage d'accord intracodeur pour le diagnostic (typique/atypique) des GMs. Au total, 86 % d'accord intracodeur a été observé pour le diagnostic des GMs, (Kappa) $k = 0.70$ [95 % IC 0.50, 0.87], $p < .05$.

Tableau 18. Pourcentage d'accord intracodeur pour le diagnostic des GMs

Codage 2	Codage 1	
	Typique	Atypique
Typique	27 %	0 %
Atypique	14 %	59 %

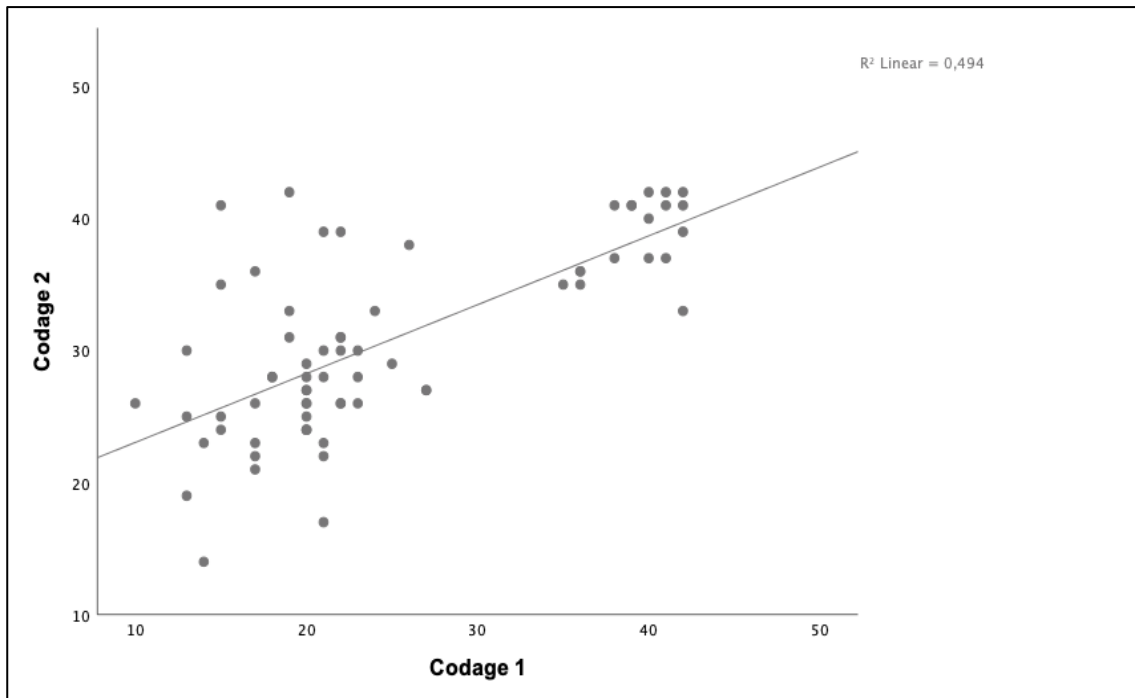
Le tableau 19 montre le pourcentage d'accord intercodeur pour le diagnostic des GMs. Au total, 77 % d'accord intercodeur a été observé pour le diagnostic des GMs, (Kappa) $k = 0.50$, [95 % IC 0.28, 0.70], $p < .05$.

Tableau 19. Pourcentage d'accord intercodeur pour le diagnostic des GMs

Codeur 2	Codeur 1	
	Typique	Atypique
Typique	22 %	19 %
Atypique	4 %	55 %

La figure 16 illustre la fiabilité intracodeur pour le score optimal des GMs. Un bon accord intracodeur a été observé en ce qui concerne le score optimal des GMs, ICC (coefficient de corrélation intra-classe) = 0.80, 95 % [0.70, 0.88], $p < .05$.

Figure 16. Fiabilité intracodeur pour le score optimal des GMs



Le tableau 20 détaille l'accord intracodeur pour les critères optimaux des GMs. L'accord intracodeur pour les critères optimaux se distribue en : bonne fiabilité pour 35 % ($n = 7$) ; fiabilité excellente pour 30 % ($n = 6$) ; fiabilité modérée pour 10 % ($n = 2$).

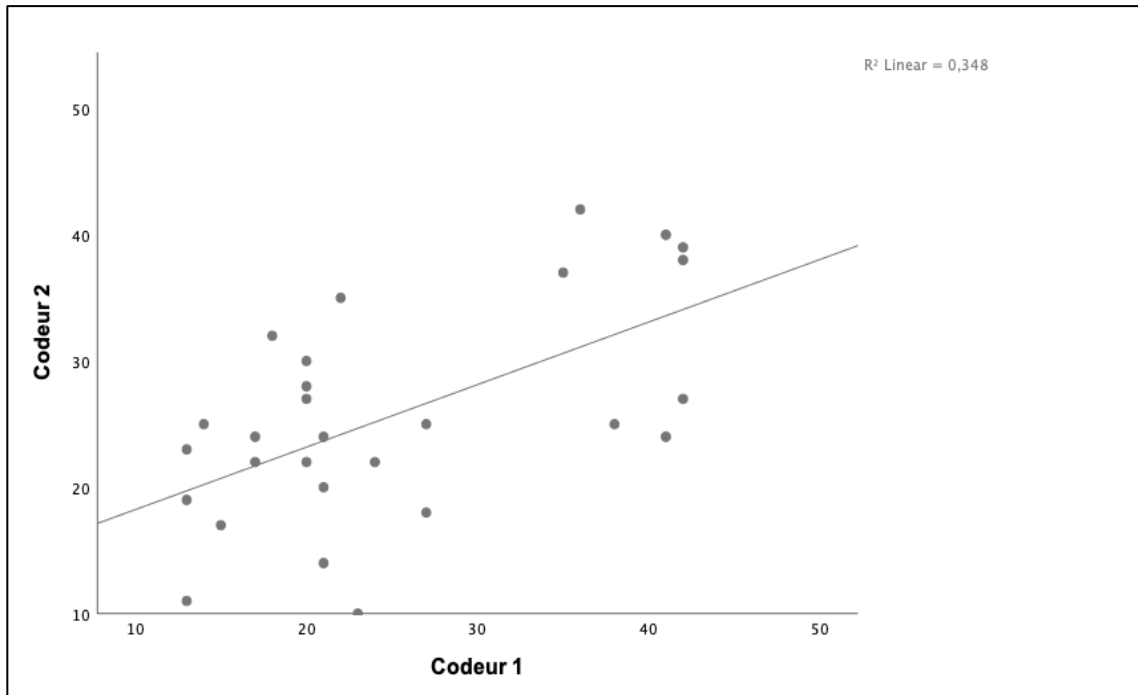
Tableau 20. Accord intracodeur pour les critères optimaux des GMs

	ICC	95 % IC	
		Inférieur	Supérieur
Cou impliqué	0.65*	0.44	0.79
Rotations du tronc	0.81*	0.70	0.88
Membres supérieurs			
Amplitude	0.62*	0.38	0.77
Vitesse	0.80*	0.70	0.88
Occupation spatiale	0.80*	0.64	0.86
Rotations proximales	0.70*	0.50	0.81
Rotations distales	0.70*	0.52	0.81
Fluctuations au début	0.10	-0.45	0.44
Fluctuations à la fin	0.11	-0.42	0.45
Tremblement	0.80*	0.69	0.88
Cramped	0.21	-0.28	0.51
Membres inférieurs			
Amplitude	0.80*	0.66	0.88
Vitesse	0.86*	0.79	0.91
Occupation spatiale	0.65*	0.44	0.80
Rotations proximales	0.66*	0.45	0.80
Rotations distales	0.66*	0.45	0.80
Fluctuations au début	0.33	-0.74	0.60
Fluctuations à la fin	0.32	-0.91	0.58
Tremblement	0.50*	0.20	0.69
Cramped	0.47*	0.13	0.68
Score optimal des GMs	0.80*	0.70	0.88

ICC, coefficient de corrélation intraclass ; IC, l'intervalle de confiance est une plage qui contient la valeur réelle sur la population ; *p < .05.

La figure 17 illustre la fiabilité intercodeur pour le score optimal des GMs. Un bon accord a été observé concernant la fiabilité intercodeur pour le score optimal des GMs, ICC = 0.73, [95 % IC 0.42, 0.87], p < .05.

Figure 17. Fiabilité intercodeur pour le score optimal des GMs



Le tableau 21 montre l'accord intercodeur pour les critères optimaux des GMs. L'accord intracodeur pour les critères optimaux se distribue en : fiabilité modérée pour 5 % (n = 1) ; bonne fiabilité pour 10 % (n = 2).

Tableau 21. Accord intercodeur pour les critères optimaux des GMs

	ICC	95 % IC	
		Inférieur	Supérieur
Cou impliqué	0.44	-0.21	0.74
Rotations du tronc	0.67*	0.26	0.84
Membres supérieurs			
Amplitude	0.36	-0.40	0.71
Vitesse	0.31	-0.50	0.69
Occupation spatiale	0.60*	0.11	0.81
Rotations proximales	0.02	-1.14	0.55
Rotations distales	-0.09	-1.40	0.50
Fluctuations au début	0.32	-0.49	0.70
Fluctuations à la fin	0.31	-0.50	0.69
Tremblement	0.12	-0.91	0.60
Cramped	0.39	-0.34	0.72
Membres inférieurs			
Amplitude	0.50	-1.09	0.56
Vitesse	0.31	-0.50	0.69
Occupation spatiale	0.37	-0.36	0.71
Rotations proximales	0.29	-0.56	0.68
Rotations distales	-0.30	-1.84	0.40
Fluctuations au début	0.54*	0.01	0.80
Fluctuations à la fin	0.48	-0.15	0.77
Tremblement	0.35	-0.41	0.70
Cramped	0.42	-0.26	0.74
Score optimal des GMs	0.73*	0.42	0.88

ICC, coefficient de corrélation intraclasse ; IC, l'intervalle de confiance est une plage qui contient la valeur réelle sur la population ; *p < .05.

7.3. Discussions

À fin de valider une base de données sur les GMs qui, en tant que mesure neuromotrice, permettra le développement d'une mesure automatisée destinée au nouveau-né prématuré, cette étude a évalué quatre propriétés psychométriques des

GMs : la validité de construit, la validité de critère, la fiabilité intracodeur et la fiabilité intercodeur. Les principaux résultats sont présentés dans les paragraphes suivants.

7.3.1. Validité

Cette étude a tout d'abord confirmé des associations significatives entre les GMs et des facteurs cliniques qui déterminent un risque neuromoteur. Ces résultats appuieraient la validité du construit des GMs pour l'enfant né prématurément. Selon le modèle multivarié, le diagnostic (typique/atypique) des GMs aussi bien que le score optimal des GMs ont été affectés par le score d'APGAR, une échographie cérébrale atypique et le diagnostic de l'HIV. Premièrement, bien que le score d'APGAR évalue des composants neurologiques (ex. le tonus musculaire), sa validité pour identifier un risque neurologique demeure conflictuelle. Alors qu'une étude ne montre pas de relation significative entre le score d'APGAR et le neurodéveloppement, une autre étude a découvert qu'un bas score d'APGAR prédit les performances neuromotrices moins bonnes (Heineman et al., 2010; Nelson & Ellenberg, 1981). Deuxièmement, les lésions cérébrales comme l'HIV, étant détectées par l'échographie cérébrale néonatale, sont associées à : des perturbations neurologiques dans des étapes précoces et des troubles neurodéveloppementaux chez l'enfant né prématurément (Anderson et al., 2017 ; George et al., 2018b ; Mukerji et al., 2015 ; Pfahl et al., 2018 ; Weisglas-Kuperus, Baerts, Fetter, & Sauer, 1992 ; Hintz et al., 2005 ; Stoll, Hansen, Adams-Chapman, et al., 2004 ; Kelly et al., 2019). Plus particulièrement, nous avons observé que le diagnostic (typique/atypique) des GMs est associé à une échographie cérébrale atypique (toutes les lésions et niveaux de gravité confondus), alors que le score des GMs est associé au diagnostic de l'HIV. Cela pourrait indiquer la validité de l'analyse quantitative des GMs pour appuyer un diagnostic différentiel.

Ensuite, cette étude démontre que le diagnostic (typique/atypique) des GMs et le score des GMs varient quant à la précision diagnostique d'une lésion cérébrale. D'une part, nos résultats ne sont pas concluants quant à la validité de critère du diagnostic (typique/atypique) des GMs par rapport à l'HIV et la LPV. Dans cette mesure, le diagnostic des GMs présente une bonne sensibilité (0.87 - 0.90) pour identifier les enfants qui présentent HIV ou LPV, mais le diagnostic des GMs présente une mauvaise spécificité (0.28 – 0.31) pour identifier les enfants n'ayant pas de HIV ou LPV. Bien que ces résultats soient confirmés dans d'autres études, il est pourtant important de les considérer avec précaution (de Vries & Bos, 2010 ; Garcia, Gherpelli, & Leone, 2004). Ainsi, une mesure qui présente une bonne sensibilité, mais une mauvaise spécificité implique un taux élevé de faux positives, c'est-à-dire des enfants présentant des GMs atypiques n'ayant pas de lésion cérébrale. Ces résultats doivent par conséquent être pris avec précaution au moment de les généraliser sur la population, car la sensibilité et la spécificité dépendent des caractéristiques de l'échantillon observé. Ainsi, ces résultats pourraient changer sur un échantillon différent (Eusebi, 2013). De plus, un trouble neuromoteur peut être également occasionné par des facteurs autres que les lésions cérébrales. D'ailleurs, le faible taux de faux négatifs (enfants présentant des GMs typiques ayant des lésions cérébrales) rend compte d'une bonne (0.94 – 0.81) VPN (valeur prédictive négative) dans la mesure où un diagnostic de GMs typiques dénote bien l'absence d'une lésion cérébrale. Cependant, ces résultats doivent aussi être interprétés avec précaution, puisque les valeurs prédictives, comme la VPN, dépendent de la prévalence de la maladie, en l'occurrence une lésion cérébrale. Par exemple, l'HIV atteint 16 % des enfants nés extrêmement prématurés et 0.5 % — 0.06 % de ceux modérément ou tardivement prématurés (Laptook, 2013). Cependant, 15 % de notre échantillon a présenté une HIV. De même, la prévalence de la LPV est 1.7 % chez les grands prématurés et moins de 0.4 % chez les modérément ou tardivement prématurés (Laptook, 2013). Cependant, 22 % de notre échantillon a présenté une LPV. D'autre part, nos résultats sont concluants quant à la validité de critère du score optimal des GMs par rapport aux lésions cérébrales, car une association significative entre le score optimal des GMs et l'HIV

ou la LPV a été démontrée. En effet, le diagnostic positif pour une lésion cérébrale entraîne la diminution du score optimal des GMs. Bien que, à notre connaissance, ce résultat ne soit pas démontré dans d'autres études, il est possible que l'association entre le score des GMs et les lésions cérébrales reflète une validité plus importante de l'approche quantitative (analyse détaillée) par rapport à l'approche qualitative (analyse globale) des GMs.

7.3.2. *Fiabilité*

La fiabilité intra/intercodeur des GMs varie concernant le diagnostic (typique/atypique) des GMs et le score optimal des GMs. En ce qui concerne le diagnostic des GMs, une fiabilité intracodeur modérée ($k = 0.70$) et une fiabilité intercodeur faible ($k = 0.50$) ont été démontrées. La plupart des études montrent une fiabilité forte ($k > 0.85$) intra/inter codeur. Cependant, ces études ont été réalisées par des tuteurs experts dans les GMs (Heineman & Hadders-Algra, 2008 ; Spittle, Doyle, & Boyd, 2008). Cependant, nos résultats ont été confirmés par d'autres études réalisées dans un cadre clinique (Bernhardt, Marbacher, Hilfiker, & Radlinger, 2011). À propos du score optimal des GMs, une bonne fiabilité intracodeur et une bonne fiabilité intercodeur ont été prouvées. Bien que ces résultats soient démontrés dans d'autres études (ICC entre 0.55 et 1.00), nous avons pourtant eu des désaccords concernant les critères optimaux des GMs (composants de l'évaluation détaillée des GMs dont la somme produit le score optimal. Voir les tableaux 20 et 21) (Yeh et al., 2016). Il a été suggéré que des facteurs subjectifs, comme la calibration de la perception visuelle, affectent la fiabilité des composants optimaux des GMs. En d'autres termes, la précision du jugement peut diminuer lorsque de nombreux GMs atypiques sont évalués dans une série sans visionner une vidéo de GMs typiques entre eux pour calibrer la perception. De même, sauf pour un cas, tous les désaccords observés dans la fiabilité intra/inter codeur, pour le diagnostic et pour le score optimal des GMs, correspondaient à des

désaccords entre la classification pauvre répertoire et la classification typique. En effet, plutôt qu'une difficulté dans la procédure de codage, les désaccords intra/inter codeur reflèteraient la nature des GMs pauvre répertoire dont la précision diagnostic est avérée d'être incertaine.

Encadré 7. Résumé du chapitre VII

Un outil automatisé requiert d'être programmé avec une mesure neuromotrice qui s'avère être validée pour le nouveau-né prématuré. D'ailleurs, la validation des GMs chez le nouveau-né prématuré étant hospitalisé reste un sujet peu exploré. Cette deuxième étude concerne la validation d'une mesure neuromotrice qui, sous forme de base de données des GMs, fut obtenue auprès de 69 nouveau-nés prématurés hospitalisés. Les objectifs de cette étude visent l'évaluation de trois propriétés psychométriques des GMs à préterme : la validité de construit par la relation des GMs avec des variables cliniques qui représentent un risque neurologique ; la validité de critère des GMs concernant le diagnostic d'une lésion cérébrale ; la fiabilité intra/intercodeur des GMs. Trois principaux résultats ont été obtenus. Premièrement, la validité de construit concernant le diagnostic (typique/atypique) des GMs et le score optimal des GMs fut confirmée par des associations avec une échographie cérébrale atypique et l'HIV. Deuxièmement, la validité de critère varie par rapport au diagnostic (typique/atypique) des GMs et le score optimal des GMs. D'une part, le diagnostic des GMs présente une bonne sensibilité (0.87 – 0.90) pour identifier les enfants ayant une HIV ou LPV, mais une mauvaise spécificité (0.28 – 0.31) pour identifier ceux n'en ayant pas. Cela comporte un taux élevé de faux positives, mais un taux faible de faux négatifs : les GMs atypiques n'impliquent pas forcément une lésion cérébrale, alors que les GMs typiques dénotent bien son absence. D'autre part, le score optimal des GMs et les lésions cérébrales présentent une relation significative, car le diagnostic de l'HIV ou la LPV entraînent la diminution du score des GMs. Cela mettrait en relief la précision d'une approche objective pour l'analyse des GMs. Troisièmement, la fiabilité intra/intercodeur des GMs varie concernant le diagnostic (typique/atypique) des GMs et le score optimal des GMs. L'accord intra/intercodeur pour le diagnostic des GMs varie entre modéré et faible, alors que l'accord intra/intercodeur pour le score des GMs est bon. Cela pourrait refléter deux aspects des GMs : la difficulté de la classification pauvre répertoire pour distinguer entre un développement typique et celui atypique ; la mise en relief de la précision de l'approche quantitative pour analyser les GMs.

Conclusions

Cette thèse rapporte la constitution et validation d'une base de données des GMs qui, en tant que mesure neuromotrice, permettra le développement d'un outil automatique destiné au nouveau-né prématuré. La base de données ainsi composée inclut 785 séquences des GMs, dont 347 correspondent à la période préterme.

Deux approches évaluatives permettent de décrire et valider cette base de données des GMs. L'approche qualitative qui dichotomise les GMs en typiques et atypiques. L'approche quantitative qui consiste à donner un score optimal aux GMs, ainsi que des sous-scores aux composants détaillés des GMs. L'approche qualitative aussi bien que l'approche quantitative présentent des avantages et des inconvénients qu'il faut considérer. Cependant, ces deux approches pourraient être utilisées d'une façon complémentaire pour évaluer la fonction neuromotrice chez le nouveau-né prématuré. En effet, ces deux approches doivent d'être considérées lors du développement d'un outil automatique destiné à évaluer les GMs chez le nouveau-né prématuré. Premièrement, l'approche qualitative ne permet pas d'identifier des changements neuromoteurs étant liés au développement. De son côté, l'approche quantitative des GMs permet de préciser des changements développementaux étant relatifs à des aspects spécifiques (ex. une implication plus importante du cou). Deuxièmement, l'approche qualitative permet de tracer l'évolution clinique de la fonction neuromotrice entre les périodes préterme, terme et post terme. Bien que l'approche quantitative ne permette pas de tracer l'évolution globale de la neuromotricité, du point de vue clinique, elle permet pourtant de préciser des améliorations cliniques sur des aspects très spécifiques (ex. une vitesse plus variable). De même, l'approche qualitative et l'approche quantitative sont réciproquement confirmées, car les GMs atypiques avaient les scores les plus bas et les GMs typiques avaient les scores les plus hauts. Troisièmement, la validité de l'approche qualitative est étayée par l'association des GMs atypiques avec une

échographie cérébrale atypique. Cependant, l'approche qualitative entraîne un taux important de faux positifs par rapport au diagnostic d'une lésion cérébrale. La validité de l'approche quantitative repose sur la diminution du score des GMs avec le diagnostic d'une HIV (hémorragie intraventriculaire) et/ou d'une LPV (leucomalacie périventriculaire). Quatrièmement, la fiabilité intra/inter codeur dans le cadre de l'approche qualitative des GMs varie entre modéré et faible ce qui représente un désaccord pour près d'un tiers des cas évalués. La fiabilité intra/inter codeur pour l'approche quantitative est meilleure. Plus particulièrement, l'accord intra codeur a été jugé bon (CCI=0.80). En effet, pour des résultats cliniquement pertinents, une fiabilité à partir de 80 % d'accord est souhaitable.

Cette étude présente des points faibles et forts. D'une part, la perte de données dans le suivi à la période terme et post terme, indiquerait la nécessité de confirmer les résultats longitudinaux. De même, comme des indices épidémiologiques peuvent être affectés par l'échantillon ou la prévalence, il pourrait être recommandé de reproduire cette étude avec un échantillon de cohorte plus large et/ou plus homogène. Cela permettrait d'obtenir des données plus précises. D'autre part, le recrutement par admission, les évaluations indépendantes, le recueil et l'évaluation systématiques des données ont permis de diminuer le risque de biais en ce qui concerne notamment la mesure et la procédure d'évaluation.

Ce travail a donné lieu à une collaboration avec l'équipe de l'Institut des Systèmes Intelligents et de Robotique (ISIR). Un premier stage de master a été réalisé pour mettre au point des tests informatisés de reconnaissance des GMs (Annexe 4). D'autres travaux restent à venir en prenant comme références notre base de données. Plus particulièrement, nos données serviront de référence pour la constitution et la validation préliminaire des outils d'analyse automatisée de la neuromotricité.

Références

- Aaronson, N., & et al. (2002). Assessing health status and quality-of-life instruments : Attributes and review criteria. *Quality of Life Research*, 193–205.
- Abily-Donval, L., Pinto-Cardoso, G., Chadie, A., Guerrot, A.-M., Torre, S., Rondeau, S., & Marret, S. (2015). Comparison in Outcomes at Two-Years of Age of Very Preterm Infants Born in 2000, 2005 and 2010. *Plos One*, 10(2), e0114567. <https://doi.org/10.1371/journal.pone.0114567>
- Adams-Chapman, I., Hansen, N., Stoll, B. J., & Higgins, R. (2008). Neurodevelopmental Outcome of Extremely Low Birth Weight Infants With Posthemorrhagic Hydrocephalus Requiring Shunt Insertion. *Pediatrics*, 121(5), e1167–e1177. <https://doi.org/10.1542/peds.2007-0423>
- Adams-Chapman, Ira, Bann, C. M., Das, A., Ronald, N., Stoll, B. J., Walsh, M. C., ... Susan, R. (2013). Neurodevelopmental Outcome of Extremely Low Birth Weight Infants with Candida Infection, 163(May 2010), 961–967. <https://doi.org/10.1016/j.jpeds.2013.04.034>. Neurodevelopmental
- Adde, L., Rygg, M., Lossius, K., Øberg, G., & Støen, R. (2007). General movement assessment: Predicting cerebral palsy in clinical practise. *Early Human Development*, 83(1), 13–18. <https://doi.org/10.1016/j.earlhumdev.2006.03.005>
- Adde, Lars, Helbostad, J. L., Jensenius, A. R., Taraldsen, G., Grunewaldt, K. H., & Støen, R. (2010). Early prediction of cerebral palsy by computer-based video analysis of general movements: a feasibility study. *Developmental Medicine & Child Neurology*, 52(8), 773–778. <https://doi.org/10.1111/j.1469-8749.2010.03629.x>
- Adde, Lars, Helbostad, J. L., Jensenius, A. R., Taraldsen, G., & Støen, R. (2009). Using computer-based video analysis in the study of fidgety movements. *Early Human Development*, 85(9), 541–547. <https://doi.org/10.1016/j.earlhumdev.2009.05.003>
- Albers S, J. G. (1994). Prognostic Significance of Spontaneous Motility in Very Immature Preterm Infants under Intensive Care Treatment. *Biol Neonate*.
- Albuquerque, P. L. De, Queiroz, M., Guerra, D. F., De, M., Lima, C., & Eickmann, S. H. (2017). Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor development in preterm infants : A comparative study with the Bayley III Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor. *Developmental Neurorehabilitation*, 00(00), 1–7. <https://doi.org/10.1080/17518423.2017.1323974>
- Allan, W. C., & Volpe, J. J. (1986). Periventricular-intraventricular hemorrhage. *Pediatric Clinics of North America*, 33(1), 47–63. <https://doi.org/10.1016/S0031->

3955(16)34969-0

- Allendorf, A., Dewitz, R., Weber, J., Bakthiar, S., Schloesser, R., & Rolle, U. (2018). Necrotizing enterocolitis as a prognostic factor for the neurodevelopmental outcome of preterm infants - match control study after 2 years. *Journal of Pediatric Surgery*, 53(8), 1573–1577. <https://doi.org/10.1016/j.jpedsurg.2018.01.006>
- Altman, D. G., & Bland, J. M. (1994). Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *Bmj*. <https://doi.org/10.1136/bmj.308.6943.1552>
- Anderson, P. J., & Doyle, L. W. (2006). Neurodevelopmental Outcome of Bronchopulmonary Dysplasia. *Seminars in Perinatology*, 30(4), 227–232. <https://doi.org/10.1053/j.semperi.2006.05.010>
- Anderson, P. J., Treyvaud, K., Neil, J. J., Cheong, J. L. Y., Hunt, R. W., Thompson, D. K., ... Inder, T. E. (2017). Associations of Newborn Brain Magnetic Resonance Imaging with Long-Term Neurodevelopmental Impairments in Very Preterm Children. *Journal of Pediatrics*, 187, 58-65.e1. <https://doi.org/10.1016/j.jpeds.2017.04.059>
- Arnaud, C., Daubisse-Marliac, L., White-Koning, M., Pierrat, V., Larroque, B., Grandjean, H., ... Kaminski, M. (2007). Prevalence and associated factors of minor neuromotor dysfunctions at age 5 years in prematurely born children: The EPIPAGE study. *Archives of Pediatrics and Adolescent Medicine*, 161(11), 1053–1061. <https://doi.org/10.1001/archpedi.161.11.1053>
- Back, S. A., Riddle, A., & McClure, M. M. (2007). Maturation-dependent vulnerability of perinatal white matter in premature birth. *Stroke*, 38(2 PART 2), 724–730. <https://doi.org/10.1161/01.STR.0000254729.27386.05>
- Baev, K. V., & Shimansky, Y. P. (1992). Principles of organization of neural systems controlling automatic movements in animals. *Progress in Neurobiology*, 39(1), 45–112. [https://doi.org/10.1016/0301-0082\(92\)90030-I](https://doi.org/10.1016/0301-0082(92)90030-I)
- Ballabh, P. (2010). Intraventricular Hemorrhage in Premature Infants: Mechanism of Disease. *Pediatric Research*, 67(1), 1–8. <https://doi.org/10.1203/PDR.0b013e3181c1b176>.Intraventricular
- Barbeau, H., & Rossignol, S. (1987). Recovery of locomotion after chronic spinalization in the adult cat. *Brain Research*, 412(1), 84–95. [https://doi.org/10.1016/0006-8993\(87\)91442-9](https://doi.org/10.1016/0006-8993(87)91442-9)
- Bashiri, A., Burstein, E., & Mazor, M. (2006). Cerebral palsy and fetal inflammatory response syndrome: A review. *Journal of Perinatal Medicine*, 34(1), 5–12. <https://doi.org/10.1515/JPM.2006.001>
- Bassler, D., Stoll, B. J., Schmidt, B., Elizabeth, V., Roberts, R. S., Robertson, C. M. T., & Reg, S. (2010). Using a Count of Neonatal Morbidities to Predict Poor Outcome in Extremely Low Birth Weight Infants: Added Role of Neonatal Infection. *Pediatrics*, 123(1), 313–318. <https://doi.org/10.1542/peds.2008-0377>.Using
- Beaino, G., Khoshnood, B., Kaminski, M., Pierrat, V., Marret, S. Ép., Matis, J., ...

- Ancel, P. Y. (2010). Predictors of cerebral palsy in very preterm infants: The EIPAGE prospective population-based cohort study. *Developmental Medicine and Child Neurology*, 52(6). <https://doi.org/10.1111/j.1469-8749.2010.03612.x>
- Bénédicte Coulm, Camille Bonnet, Béatrice Blondel, Alexis Vanhaesebrouck, Annick Vilain, Jeanne Fresson, S. R. (2017). Enquête nationale périnatale Rapport 2016.
- Berge, P., Adde, L., Espinosa, G., & Stavdahl, O. (2008). ENIGMA – Enhanced interactive general movement assessment. *Expert Systems with Applications*, 34(4), 2664–2672. <https://doi.org/10.1016/j.eswa.2007.05.024>
- Berkelhamer, S. K., Mestan, K. K., & Steinhorn, R. (2018). An update on the diagnosis and management of bronchopulmonary dysplasia (BPD)-associated pulmonary hypertension. *Seminars in Perinatology*. <https://doi.org/10.1053/j.semperi.2018.09.005>
- Bernhardt, I., Marbacher, M., Hilfiker, R., & Radlinger, L. (2011). Inter- and intra-observer agreement of Prechtl's method on the qualitative assessment of general movements in preterm, term and young infants. *Early Human Development*, 87(9), 633–639. <https://doi.org/10.1016/j.earlhumdev.2011.04.017>
- Bewick, V., Cheek, L., & Ball, J. (2003). Statistics review 7: Correlation and regression. *Critical Care (London, England)*, 7(6), 451–459. <https://doi.org/10.1186/cc2401>
- Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 8: Qualitative data - tests of association. *Critical Care (London, England)*, 8(1), 46–53. <https://doi.org/10.1186/cc2428>
- Bhandari, A., & Panitch, H. (2018). An update on the post-NICU discharge management of bronchopulmonary dysplasia. *Seminars in Perinatology*. <https://doi.org/10.1053/j.semperi.2018.09.011>
- Blauw-Hospers, C. H., de Graaf-Peters, V. B., Dirks, T., Bos, a. F., & Hadders-Algra, M. (2007). Does early intervention in infants at high risk for a developmental motor disorder improve motor and cognitive development? *Neuroscience and Biobehavioral Reviews*, 31(8), 1201–1212. <https://doi.org/10.1016/j.neubiorev.2007.04.010>
- Blauw-Hospers, Cornill H, & Hadders-Algra, M. (2005). A systematic review of the effects of early intervention on motor development. *Developmental Medicine & Child Neurology*, 47(6), 421–432. <https://doi.org/10.1111/j.1469-8749.2005.tb01165.x>
- Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A.-B., ... Lawn, J. (2013). Born too soon: the global epidemiology of 15 million preterm births. *Reproductive Health*, 10 Suppl 1(Suppl 1), S2. <https://doi.org/10.1186/1742-4755-10-S1-S2>
- Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A. B., Narwal, R., ... Lawn, J. E. (2012). National, regional, and worldwide estimates of preterm

- birth rates in the year 2010 with time trends since 1990 for selected countries: A systematic analysis and implications. *The Lancet*, 379(9832), 2162–2172. [https://doi.org/10.1016/S0140-6736\(12\)60820-4](https://doi.org/10.1016/S0140-6736(12)60820-4)
- Blencowe, H., Lee, A. C., Cousens, S., Bahalim, A., Narwal, R., Zhong, N., ... Lawn, J. E. (2013). Preterm birth–associated neurodevelopmental impairment estimates at regional and global levels for 2010. *Pediatric Research*, 74(december), 17–34. <https://doi.org/10.1038/pr.2013.204>
- Blencowe, H., Vos, T., Lee, A. C., Philips, R., Lozano, R., Alvarado, M. R., ... Lawn, J. E. (2013). Estimates of neonatal morbidities and disabilities at regional and global levels for 2010: introduction, methods overview, and relevant findings from the Global Burden of Disease study. *Pediatric Research*, 74, 4–16. <https://doi.org/10.1038/pr.2013.203>
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, 111(4), 1061–1071. <https://doi.org/10.1037/0033-295X.111.4.1061>
- Bos, a F., van Asperen, R. M., de Leeuw, D. M., & Prechtl, H. F. (1997). The influence of septicaemia on spontaneous motility in preterm infants. *Early Hum Dev*, 50(1), 61–70. <https://doi.org/S0378378297000935> [pii]
- Bos, A. F., Van Braeckel, K. N. J. A., Hitzert, M. M., Tanis, J. C., & Roze, E. (2013). Development of fine motor skills in preterm infants. *Developmental Medicine and Child Neurology*, 55(SUPPL.4), 1–4. <https://doi.org/10.1111/dmcn.12297>
- Bosanquet, M., Copeland, L., Ware, R., & Boyd, R. (2013). A systematic review of tests to predict cerebral palsy in young children. *Developmental Medicine & Child Neurology*, 55(5), 418–426. <https://doi.org/10.1111/dmcn.12140>
- Bose, C., Kuban, K., Laughon, M., Ehrenkranz, R. A., Van Marter, L. J., O’Shea, M. T., ... Leviton, A. (2009). Chronic Lung Disease and Developmental Delay at 2 Years of Age in Children Born Before 28 Weeks’ Gestation. *Pediatrics*, 124(2), 637–648. <https://doi.org/10.1542/peds.2008-2874>
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... De Vet, H. C. W. (2007). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Veterinary Clinical Pathology*, 36(1), 8–12. <https://doi.org/10.1111/j.1939-165X.2007.tb00175.x>
- Brown, N., & Spittle, A. (2014). Neurobehavioral evaluation in the preterm and term infant. *Current Pediatric Reviews*, 10(1), 65–72. <https://doi.org/10.2174/157339631001140408121310>
- Bruggink, J. L. M., Koenraad, N., Braeckel, V., & Bos, A. F. (2010). The Early Motor Repertoire of Children Born Preterm Is Associated With Intelligence at School Age. <https://doi.org/10.1542/peds.2009-2117>
- Bruton, A, Conway, J H and Holgate, S. T. (2000). Reliability: What is it, and how is it measured? *Physiotherapy*, 86(2), 94–99. [https://doi.org/10.1016/S0031-9406\(05\)61211-4](https://doi.org/10.1016/S0031-9406(05)61211-4)
- Bulpitt, C. J. (1987). Confidence Intervals. *The Lancet*, 329(8531), 494–497.

[https://doi.org/10.1016/S0140-6736\(87\)92100-3](https://doi.org/10.1016/S0140-6736(87)92100-3)

- Bussel, B., Roby-Brami, A., Azouvi, A., Biraben, A., Yakovleff, A., & Held, J. (1988). Myoclonus in a patient with spinal cord transection. *Brain*, *111*(5), 1235–1245. <https://doi.org/10.1093/brain/111.5.1235>
- Butcher, P. R., van Braeckel, K., Bouma, A., Einspieler, C., Stremmelaar, E. F., & Bos, A. F. (2009). The quality of preterm infants' spontaneous movements: an early indicator of intelligence and behaviour at school age. *Journal of Child Psychology and Psychiatry*, *50*(8), 920–930. <https://doi.org/10.1111/j.1469-7610.2009.02066.x>
- Calancie, B. (2006). Spinal myoclonus after spinal cord injury. *Journal of Spinal Cord Medicine*, *29*(4), 413–424. <https://doi.org/10.1080/10790268.2006.11753891>
- Campbell, S. K., Kolobe, T. H. A., Osten, E. T., Lenke, M., & Girolami, G. L. (1995). Construct validity of the test of infant motor performance. *Physical Therapy*, *75*(7), 585–596. <https://doi.org/10.1093/ptj/75.7.585>
- Campbell, S., & Kolobe, T. (2000). Concurrent Validity of the Test of Infant Motor Performance with the Alberta Infant Motor Scale.
- Campbell, S K, Swanlund, A., Smith, E., Liao, P., & Zawacki, L. (2008). Validity of the TIMPSI for estimating concurrent. *Pediatric Physical Therapy*, *20*(1), 3–10. <https://doi.org/10.1097/PEP.0b013e31815f66a6.VALIDITY>
- Campbell, Suzann K., & Hedeker, D. (2001). Validity of the Test of Infant Motor Performance for discriminating among infants with varying risk for poor motor outcome. *Journal of Pediatrics*, *139*, 546–551. <https://doi.org/10.1067/mpd.2001.117581>
- Carrington M, Lachs M, Feinstein A. (1995). Use of Methodological Standards in Diagnostic Test Research. *JAMA*, *274*, 645–651.
- Carrington, T. (1997). How to read a paper. Papers that report diagnostic or screening tests. *BMJ (Clinical Research Ed.)*, *315*(7107), 540–543. <https://doi.org/10.1136/bmj.315.7107.540>
- Case-Smith, J. (1996). Analysis of Current Motor Development Theory and Recently Published Infant Motor Assessments.
- Casey, B. M., McIntire, D. D., & Leveno, K. J. (2001). The Continuing Value of the Apgar Score for the Assessment of Newborn Infants. *Obstetric and Gynecologic Survey*, *56*(7), 406–407. <https://doi.org/10.1097/00006254-200107000-00010>
- Castaño, A., & Maurer, M. S. (2015). HHS Public Access, *20*(2), 163–178. <https://doi.org/10.1007/s10741-014-9462-7.Natural>
- Chabra, S. (2018). Small For Gestational Age and Low Birthweight: Distinct Entities. *The American Journal of Gastroenterology*, *113*(3), 441–442. <https://doi.org/10.1038/ajg.2017.499>
- Chan, E., Leong, P., Malouf, R., & Quigley, M. A. (2016). Long-term cognitive and school outcomes of late-preterm and early-term births: A systematic review. *Child: Care, Health and Development*, *42*(3), 297–312.

<https://doi.org/10.1111/cch.12320>

- Chen, G., Faris, P., Hemmelgarn, B., Walker, R. L., & Quan, H. (2009). Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Medical Research Methodology*, 9(1), 5. <https://doi.org/10.1186/1471-2288-9-5>
- Chen, L., leee, M., Hoey, J., Nugent, C., leee, M., Cook, D., & Yu, Z. (2012). Sensor-based Activity Recognition. *Syst Man Cybern.*
- Cheong, J. L., Doyle, L. W., Burnett, A. C., Lee, K. J., Walsh, J. M., Potter, C. R., ... Spittle, A. J. (2017). Association between moderate and late preterm birth and neurodevelopment and social-emotional development at age 2 years. *JAMA Pediatrics*, 171(4), 1–7. <https://doi.org/10.1001/jamapediatrics.2016.4805>
- Cheong, J. L. Y., & Doyle, L. W. (2018). An update on pulmonary and neurodevelopmental outcomes of bronchopulmonary dysplasia. *Seminars in Perinatology*. <https://doi.org/10.1053/j.semperi.2018.09.013>
- Chiriboga, C. A., Kuhn, L., Durkin, M., Bellinger, D., K, K. C., & Sanocka, U. (2010). Factors associated with microcephaly at school age in a very-low-birthweight population. *Developmental Medicine & Child Neurology*, 45(12), 796–801. <https://doi.org/10.1111/j.1469-8749.2003.tb00893.x>
- Cicchetti, D. V. (1994). Interreliability Standards in Psychological Evaluations. *Psychological Assessment*, (4), 284–290.
- Cioni, G., Ferrari, F., Einspieler, C., Paolicelli, P., Barbani, T., & Prechtl, H. (1997). Comparison between observation of spontaneous movements and neurologic examination in preterm infants. *The Journal of Pediatrics*, 130(5), 704–711. [https://doi.org/10.1016/S0022-3476\(97\)80010-8](https://doi.org/10.1016/S0022-3476(97)80010-8)
- Cioni, G., & Prechtl, H. (1990). Preterm and early postterm motor behaviour in low-risk premature infants. *Early Human Development*, 23(3), 159–191. [https://doi.org/10.1016/0378-3782\(90\)90012-8](https://doi.org/10.1016/0378-3782(90)90012-8)
- Cohen, A. H., Guan, L., Harris, J., Jung, R., & Kiemel, T. (1996). Interaction between the caudal brainstem and the lamprey central pattern generator for locomotion. *Neuroscience*, 74(4), 1161–1173. [https://doi.org/10.1016/S0306-4522\(96\)00184-4](https://doi.org/10.1016/S0306-4522(96)00184-4)
- Cohen, J. (1960). A coefficient of Agreement for nominal scales, XX(1), 37–46.
- Constantinou, J. C., Adamson-Macedo, E. N., Mirmiran, M., & Fleisher, B. E. (2007). Movement, imaging and neurobehavioral assessment as predictors of cerebral palsy in preterm infants. *Journal of Perinatology*, 27(4), 225–229. <https://doi.org/10.1038/sj.jp.7211664>
- Craciunoiu, O., & Holsti, L. (2016). A Systematic Review of the Predictive Validity of Neurobehavioral Assessments During the Preterm Period. *Physical & Occupational Therapy In Pediatrics*, 2638(June), 1–16. <https://doi.org/10.1080/01942638.2016.1185501>
- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests.

Psychological Bulletin, (4), 281–302.

- Dalton, T. C. (2005). Arnold Gesell and the maturation controversy. *Integrative Physiological and Behavioral Science*, 40(4), 182–204. <https://doi.org/10.1007/BF02915215>
- Dammann, O., Kuban, K. C. K., & Leviton, A. (2002). Perinatal infection, fetal inflammatory response, white matter damage, and cognitive limitations in children born preterm. *Mental Retardation and Developmental Disabilities Research Reviews*, 8(1), 46–50. <https://doi.org/10.1002/mrdd.10005>
- Danner, S. M., Hofstoetter, U. S., Freundl, B., Binder, H., Mayr, W., Rattay, F., & Minassian, K. (2015). Human spinal locomotor control is based on flexibly organized burst generators. *Brain*, 138(3), 577–588. <https://doi.org/10.1093/brain/awu372>
- Darsaklis, V. (2011). Presentation de données pour meta-analyses. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Darsaklis, V., Snider, L. M., Majnemer, A., & Mazer, B. (2011). Predictive validity of Prechtl's Method on the Qualitative Assessment of General Movements: a systematic review of the evidence. *Developmental Medicine & Child Neurology*, 53(10), 896–906. <https://doi.org/10.1111/j.1469-8749.2011.04017.x>
- Day, C. L., & Ryan, R. M. (2017). Bronchopulmonary dysplasia: New becomes old again! *Pediatric Research*, 81(1–2), 210–213. <https://doi.org/10.1038/pr.2016.201>
- de Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59(10), 1033–1039. <https://doi.org/10.1016/j.jclinepi.2005.10.015>
- de Vries, J. I., Visser, G. H., & Prechtl, H. F. (1982). The emergence of fetal behaviour. I. Qualitative aspects. *Early Human Development*, 7(4), 301–322. [https://doi.org/10.1016/0378-3782\(82\)90033-0](https://doi.org/10.1016/0378-3782(82)90033-0)
- de Vries, L. S. (1996). Neurological assessment of the preterm infant. *Acta Paediatrica (Oslo, Norway: 1992)*, 85(7), 765–771. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8819539>
- de Vries, Linda S., Eken, P., & Dubowitz, L. M. S. (1992). The spectrum of leukomalacia using cranial ultrasound. *Behavioural Brain Research*, 49(1), 1–6. [https://doi.org/10.1016/S0166-4328\(05\)80189-5](https://doi.org/10.1016/S0166-4328(05)80189-5)
- de Vries, Linda S., Van Haastert, I. L. C., Rademaker, K. J., Koopman, C., & Groenendaal, F. (2004). Ultrasound abnormalities preceding cerebral palsy in high-risk preterm infants. *Journal of Pediatrics*, 144(6), 815–820. <https://doi.org/10.1016/j.jpeds.2004.03.034>
- de Vries, N., & Bos, A. (2011). The motor repertoire of extremely low-birthweight infants at term in relation to their neurological outcome. *Developmental Medicine & Child Neurology*, 53(10), 933–937. <https://doi.org/10.1111/j.1469-8749.2011.04045.x>

- de Vries, N. K. S., & Bos, A. F. (2010). The quality of general movements in the first ten days of life in preterm infants. *Early Human Development*, *86*(4), 225–229. <https://doi.org/10.1016/j.earlhumdev.2010.03.004>
- de Vries, N. K. S., Erwich, J. J. H. M., & Bos, A. F. (2008). General movements in the first fourteen days of life in extremely low birth weight (ELBW) infants. *Early Human Development*, *84*(11), 763–768. <https://doi.org/10.1016/j.earlhumdev.2008.05.003>
- Dietz, V. (2003). Spinal cord pattern generators for locomotion. *Clinical Neurophysiology*, *114*(8), 1379–1389. [https://doi.org/10.1016/S1388-2457\(03\)00120-2](https://doi.org/10.1016/S1388-2457(03)00120-2)
- Dietz, V., Colombo, G., Jensen, L., & Baumgartner, L. (1995). Locomotor Capacity of Spinal Cord in Paraplegic Patients Materials and Methods General Procedures and Recording Methods.
- Dimitrijevic, M. R., Gerasimenko, Y., & Pinter, M. M. (1998). Evidence for a Spinal Central Pattern Generator in Humans. *Annals of the New York Academy of Sciences*, *860*(1), 360–376. <https://doi.org/10.1111/j.1749-6632.1998.tb09062.x>
- Divyen, S., Doyle, L., Anderson, P., Bear, M., Daley, A., Hunt, R., & Inder, T. (2008). Necrotizing Enterocolitis is Mediated by White Matter Abnormalities on. *The Journal of Pediatrics*, *153*(2), 170-175.e1. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0022347608001091>
- Dong, Y., & Yu, J. L. (2011). An overview of morbidity, mortality and long-term outcome of late preterm birth. *World Journal of Pediatrics*, *7*(3), 199–204. <https://doi.org/10.1007/s12519-011-0290-8>
- Doyle, L. W., Roberts, G., & Anderson, P. J. (2010). Outcomes at Age 2 Years of Infants < 28 Weeks' Gestational Age Born in Victoria in 2005. *The Journal of Pediatrics*, *156*(1), 49-53.e1. <https://doi.org/10.1016/j.jpeds.2009.07.013>
- Draper, E. S., Zeitlin, J., Fenton, A. C., Weber, T., Gerrits, J., Martens, G., ... Breart, G. (2009). Investigating the variations in survival rates for very preterm infants in 10 European regions: The MOSAIC birth cohort. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *94*(3). <https://doi.org/10.1136/adc.2008.141531>
- Dubowitz, L., Ricci, D., & Mercuri, E. (2005). The Dubowitz neurological examination of the full-term newborn. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(1), 52–60. <https://doi.org/10.1002/mrdd.20048>
- Eaton, S., Rees, C. M., & Hall, N. J. (2017). Current Research on the Epidemiology, Pathogenesis, and Management of Necrotizing Enterocolitis. *Neonatology*, *111*(4), 423–430. <https://doi.org/10.1159/000458462>
- Eber, E., & Zach, M. S. (2001). Paediatric origins of adult lung disease bullet 8: Long term sequelae of bronchopulmonary dysplasia (chronic lung disease of infancy). *Thorax*, *56*(4), 317–323. <https://doi.org/10.1136/thorax.56.4.317>
- Ecury-Goossen, G. M., Camfferman, F. A., Leijser, L. M., Govaert, P., & Dudink, J.

- (2015). State of the Art Cranial Ultrasound Imaging in Neonates. *Journal of Visualized Experiments*, (96), 1–10. <https://doi.org/10.3791/52238>
- Ehrenkranz, R., Walsh, M., Vohr, B., Jobe, A., Wright, L., Fanaroff, A., ... Poole, K. (2005). Validation of the National Institutes of Health Consensus Definition of Bronchopulmonary Dysplasia. *Pediatrics*, 116(6), 1353–1360. <https://doi.org/10.1542/peds.2005-0249>
- Einspieler, C. (2008). Early markers for unilateral spastic cerebral palsy in premature infants. *Nature Clinical Practice Neurology*, 4(4), 186–187. <https://doi.org/10.1038/ncpneuro0745>
- Einspieler, C, Bos, A., Libertus, M., & Marschik, P. (2016). The General Movement Assessment Helps Us to Identify Preterm Infants at Risk for Cognitive Dysfunction. *Frontiers in Psychology*, 7(March), 406. <https://doi.org/10.3389/fpsyg.2016.00406>
- Einspieler, C, Prayer, D., & Prechtl, H. (2012). *Fetal Behaviour: A neurodevelopmental approach*. *Journal of Chemical Information and Modeling* (Vol. 53). <https://doi.org/10.1017/CBO9781107415324.004>
- Einspieler, C, & Prechtl, H. F. R. (2005). Prechtl's assessment of general movements: A diagnostic tool for the functional assessment of the young nervous system. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 61–67. <https://doi.org/10.1002/mrdd.20051>
- Einspieler, Christa, Marschik, P. B., Pansy, J., Scheuchenegger, A., Kriebler, M., Yang, H., ... Bos, A. F. (2015). The general movement optimality score: a detailed assessment of general movements during preterm and term age. *Developmental Medicine & Child Neurology*, n/a-n/a. <https://doi.org/10.1111/dmcn.12923>
- Einspieler, Christa, Prechtl, H. F. R., Ferrari, F., Cioni, G., & Bos, A. F. (1997). The qualitative assessment of general movements in preterm, term and young infants — review of the methodology. *Early Human Development*, 50(1), 47–60. [https://doi.org/10.1016/S0378-3782\(97\)00092-3](https://doi.org/10.1016/S0378-3782(97)00092-3)
- Einspieler, C., Prechtl, H. F. R., Bos, A., Ferrari, F., & Cioni, G. (2004). Prechtl's method on the qualitative assessment of general movements in preterm, term and young infants. *Clin. Dev. Med*, 167, 1–91.
- Elie, C., & Colombet, I. (2011). Méthodes d'estimation de la reproductibilité. *Sang Thrombose Vaisseaux*, 23(3), 138–145. <https://doi.org/10.1684/stv.2011.0592>
- Eusebi, P. (2013). Diagnostic Accuracy Measures. *Cerebrovasc Dis*, 267–272. <https://doi.org/10.1159/000353863>
- Eyre, J., Miller, S., Clowry, G., Conway, E., & Watts, C. (2000). Functional corticospinal projections are established prenatally in the human foetus permitting involvement in the development of spinal motor centres. *Brain*, 123(1), 51–64. <https://doi.org/10.1093/brain/123.1.51>
- Fawke, J. (2007). Neurological outcomes following preterm birth. *Seminars in Fetal and Neonatal Medicine*, 12(5), 374–382.

<https://doi.org/10.1016/j.siny.2007.06.002>

- Feller, M. B. (1999). Spontaneous Correlated Activity in Developing Neural Circuits. *Neuron*, 22, 653–656. <https://doi.org/10.1038/nrn2759>.Mechanisms
- Fermanian, J. (2005). Validation des échelles d'évaluation en médecine physique et de réadaptation: Comment apprécier correctement leurs qualités psychométriques. *Annales de Readaptation et de Medecine Physique*, 48(6), 281–287. <https://doi.org/10.1016/j.annrmp.2005.04.004>
- Ferrari, F, Cioni, G., & Prechtl, H. F. (1990). Qualitative changes of general movements in preterm infants with brain lesions. *Early Human Development*, 23(3), 193–231. [https://doi.org/10.1016/0378-3782\(90\)90013-9](https://doi.org/10.1016/0378-3782(90)90013-9)
- Ferrari, Fabrizio, Cioni, G., Einspieler, C., Roversi, M. F., Bos, A. F., Paolicelli, P. B., ... Prechtl, H. F. R. (2002). Cramped synchronized general movements in preterm infants as an early marker for cerebral palsy. *Archives of Pediatrics & Adolescent Medicine*, 156(5), 460–467. <https://doi.org/poa10323> [pii]
- Ferrari, Fabrizio, Gallo, C., Pugliese, M., Guidotti, I., Gavioli, S., Coccolini, E., ... Bertocelli, N. (2012). Preterm birth and developmental problems in the preschool age. Part I: minor motor problems. *Journal of Maternal-Fetal and Neonatal Medicine*, 25(11), 2154–2159. <https://doi.org/10.3109/14767058.2012.696164>
- Field, A., Miles, J., & Zoe FIELD, F. (2012). *Discovering Using R Statistics*.
- Fitzgibbons, S. C., Ching, Y., Yu, D., Carpenter, J., Kenny, M., Weldon, C., ... Jaksic, T. (2009). Mortality of necrotizing enterocolitis expressed by birth weight categories. *Journal of Pediatric Surgery*, 44(6), 1072–1076. <https://doi.org/10.1016/j.jpedsurg.2009.02.013>
- Fjørtoft, T., Einspieler, C., Adde, L., & Strand, L. I. (2009). Inter-observer reliability of the “Assessment of Motor Repertoire — 3 to 5 Months” based on video recordings of infants. *Early Human Development*, 85(5), 297–302. <https://doi.org/10.1016/j.earlhumdev.2008.12.001>
- Fjørtoft, T., Grunewaldt, K. H., Løhaugen, G. C. C., Mørkved, S., Skranes, J., & Evensen, K. A. I. (2013). Assessment of motor behaviour in high-risk-infants at 3months predicts motor and cognitive outcomes in 10years old children. *Early Human Development*, 89(10), 787–793. <https://doi.org/10.1016/j.earlhumdev.2013.06.007>
- Fleiss, J. L. (1986). Confidence intervals vs significance tests: quantitative interpretation. *American Journal of Public Health*, 76(5), 587–588.
- Forsberg, H., Grillner, S., Halbertsma, J., & Rossignol, S. (1980). The locomotion of the low spinal cat. II. Interlimb coordination. *Acta Physiologica Scandinavica*, 108(3), 283–295. <https://doi.org/10.1111/j.1748-1716.1980.tb06534.x>
- Frigon, A. (2012). Central pattern generators of the mammalian spinal cord. *Neuroscientist*, 18(1), 56–69. <https://doi.org/10.1177/1073858410396101>
- Gagnon, K., Cannon, S., & Weatherstone, K. B. (2012). The Premie-Neuro:

- Opportunities and challenges for standardized neurologic assessment of the preterm infant. *Advances in Neonatal Care*, 12(5), 310–317. <https://doi.org/10.1097/ANC.0b013e318265b3fa>
- Galinsky, R., Lear, C. A., Dean, J. M., Wassink, G., Dhillon, S. K., Fraser, M., ... Gunn, A. J. (2018). Complex interactions between hypoxia-ischemia and inflammation in preterm brain injury. *Developmental Medicine and Child Neurology*, 60(2), 126–133. <https://doi.org/10.1111/dmcn.13629>
- Ganatra, H. A., Stoll, B. J., & Zaidi, A. K. M. (2010). International perspective on early-onset neonatal sepsis. *Clinics in Perinatology*, 37(2), 501–523. <https://doi.org/10.1016/j.clp.2010.02.004>
- Garcia, J. M., GherPELLI, J. L. D., & Leone, C. R. (2004). The role of spontaneous general movement assessment in the neurological outcome of cerebral lesions in preterm infants. *Jornal de Pediatria*, 80(4), 296–304. <https://doi.org/10.2223/JPED.1203>
- Gardner, M. J., & Altman, D. G. (1986). Statistics in Medicine Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal (Clinical Research Ed.)*, 292(6522), 746–750. <https://doi.org/10.1136/bmj.292.6522.746>
- Garg, M., Kurzner, S. I., Bautista, D. B., & Keens, T. G. (1988). Clinically unsuspected hypoxia during sleep and feeding in infants with bronchopulmonary dysplasia. *Pediatrics*, 81(5), 635–642. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3357725>
- George, J. M., Fiori, S., Fripp, J., Pannek, K., Guzzetta, A., David, M., ... Boyd, R. N. (2018a). Relationship between very early brain structure and neuromotor, neurological and neurobehavioral function in infants born <31 weeks gestational age. *Early Human Development*, 117(June), 74–82. <https://doi.org/10.1016/j.earlhumdev.2017.12.014>
- George, J. M., Fiori, S., Fripp, J., Pannek, K., Guzzetta, A., David, M., ... Boyd, R. N. (2018b). Relationship between very early brain structure and neuromotor, neurological and neurobehavioral function in infants born <31 weeks gestational age. *Early Human Development*, 117(November 2017), 74–82. <https://doi.org/10.1016/j.earlhumdev.2017.12.014>
- Gill, J. V., & Boyle, E. M. (2017). Outcomes of infants born near term. *Archives of Disease in Childhood*, 102(2), 194–198. <https://doi.org/10.1136/archdischild-2015-309584>
- Gosselin, J., Gahagan, S., & Amiel-Tison, C. (2005). The Amiel-Tison Neurological Assessment at Term: Conceptual and methodological continuity in the course of follow-up. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 34–51. <https://doi.org/10.1002/mrdd.20049>
- Gravem, D., Singh, M., Chen, C., Rich, J., Vaughan, J., Goldberg, K., ... Patterson, D. (2012). Assessment of Infant Movement With a Compact Wireless Accelerometer System. *Journal of Medical Devices*, 6(2), 021013. <https://doi.org/10.1115/1.4006129>

- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350. <https://doi.org/10.1007/s10654-016-0149-3>
- Gregory, K., DeForge, C., Natale, K., Phillips, M., & Van Marter, L. (2011). Necrotizing Enterocolitis in the Premature Infant. *Adv Neonatal Care.*, *11*(3), 155–166. <https://doi.org/10.1148/83.5.879>
- Grillner, & Wallén. (1985). Central pattern generators for locomotion, with special reference to vertebrate. *Ann. Rev. Neuro.Lci.*
- Grillner, & Zangger. (1984). The effect of dorsal root transection on the efferent motor pattern in the cat's hindlimb during locomotion. *Acta Physiologica Scandinavica*, *120*(3), 393–405. <https://doi.org/10.1111/j.1748-1716.1984.tb07400.x>
- Grunewaldt, K. H., Fjørtoft, T., Bjuland, K. J., Brubakk, A.-M., Eikenes, L., Håberg, A. K., ... Skranes, J. (2014). Follow-up at age 10years in ELBW children — Functional outcome, brain morphology and results from motor assessments in infancy. *Early Human Development*, *90*(10), 571–578. <https://doi.org/10.1016/j.earlhumdev.2014.07.005>
- Guillet, R., Stoll, B., Cotten, M., Gantz, M., McDonald, S., Poole, K., & Phelps, D. (2006). Association of H2-Blocker Therapy and Higher Incidence of Necrotizing Enterocolitis in Very Low Birth Weight Infants. *Pediatrics*, *117*(2), e137–e142. <https://doi.org/10.1542/peds.2005-1543>
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, *2*(10), 685–694. <https://doi.org/10.1038/35094500>
- Guzzetta, A., Belmonti, V., Battini, R., Boldrini, A., Bruna Paolicelli, P., & Cioni, G. (2007). Does the assessment of general movements without video observation reliably predict neurological outcome? *European Journal of Paediatric Neurology*, *11*(6), 362–367. <https://doi.org/10.1016/j.ejpn.2007.03.003>
- Hadchouel, A., & Delacourt, C. (2013). Premature infants bronchopulmonary dysplasia: Past and present. *Revue de Pneumologie Clinique*, *69*(4), 207–216. <https://doi.org/10.1016/j.pneumo.2013.05.003>
- Hadders-Algra, M. (2014). Early Diagnosis and Early Intervention in Cerebral Palsy. *Frontiers in Neurology*, *5*(September), 1–13. <https://doi.org/10.3389/fneur.2014.00185>
- Hadders-Algra, M., & Prechtl, H. F. R. (1992). Developmental course of general movements in early infancy. I. Descriptive analysis of change in form. *Early Human Development*, *28*(3), 201–213. [https://doi.org/10.1016/0378-3782\(92\)90167-F](https://doi.org/10.1016/0378-3782(92)90167-F)
- Hadders-Algra, M. (1993). General movements in early infancy: what do they tell us about the nervous system? *Early Human Development*, *34*(1–2), 29–37. [https://doi.org/10.1016/0378-3782\(93\)90038-V](https://doi.org/10.1016/0378-3782(93)90038-V)

- Hadders-Algra, Mijna. (2002). Two distinct forms of minor neurological dysfunction: perspectives emerging from a review of data of the Groningen Perinatal Project. *Developmental Medicine and Child Neurology*, 44(8), 561–571. <https://doi.org/10.1017/S0012162201002560>
- Hadders-Algra, Mijna. (2010). Variation and variability: key words in human motor development. *Physical Therapy*, 90(12), 1823–1837. <https://doi.org/10.2522/ptj.20100006>
- Hadders-Algra, Mijna. (2018a). Early human motor development: From variation to the ability to vary and adapt. *Neuroscience and Biobehavioral Reviews*, 90, 411–427. <https://doi.org/10.1016/j.neubiorev.2018.05.009>
- Hadders-Algra, Mijna. (2018b). Early human motor development: From variation to the ability to vary and adapt. *Neuroscience and Biobehavioral Reviews*, 90(2010), 411–427. <https://doi.org/10.1016/j.neubiorev.2018.05.009>
- Hadders-Algra, Mijna, Klip-van den Nieuwendijk, A. W., Martijn, A., & van Eyken, L. A. (1997). Assessment of general movements : towards a better evaluate brain function in young infants. *Developmental Medicine and Child Neurology*, 39, 89–99.
- Hall, N., Peters, M., Hiorns, M., Eaton, S., Pierro, A., Khoo, A., & Tighe, H. (2008). Is Necrotizing Enterocolitis Associated with Development or Progression of Intraventricular Hemorrhage? *American Journal of Perinatology*, 26(02), 139–143. <https://doi.org/10.1055/s-0028-1095177>
- Harris, S. R., Haley, S. M., Tada, W. L., & Swanson, M. W. (1984). Reliability of Observational Measures of the Movement Assessment of Infants, 64(4), 471–475.
- Hayat, T. T. A., & Rutherford, M. A. (2018). Neuroimaging perspectives on fetal motor behavior. *Neuroscience and Biobehavioral Reviews*, 92, 390–401. <https://doi.org/10.1016/j.neubiorev.2018.06.001>
- Heemskerk, A. M., Dudink, J., Plaisier, A., Govaert, P., Feijen-Roon, M., Lequin, M. H., ... van der Starre, C. (2012). Safety of routine early MRI in preterm infants. *Pediatric Radiology*, 42(10), 1205–1211. <https://doi.org/10.1007/s00247-012-2426-y>
- Heineman, K. R., Bos, A. F., & Hadders-Algra, M. (2008). The Infant Motor Profile: a standardized and qualitative method to assess motor behaviour in infancy. *Developmental Medicine & Child Neurology*, 50(4), 275–282. <https://doi.org/10.1111/j.1469-8749.2008.02035.x>
- Heineman, K. R., & Hadders-Algra, M. (2008a). Evaluation of neuromotor function in infancy-a systematic review of available methods. *Journal of Developmental and Behavioral Pediatrics*, 29(4), 315–323. <https://doi.org/10.1097/DBP.0b013e318182a4ea>
- Heineman, K. R., & Hadders-Algra, M. (2008b). Evaluation of Neuromotor Function in Infancy-A Systematic Review of Available Methods. *Journal of Developmental & Behavioral Pediatrics*, 29(4), 315–323.

<https://doi.org/10.1097/DBP.0b013e318182a4ea>

- Heineman, K. R., La Bastide-Van Gemert, S., Fidler, V., Middelburg, K. J., Bos, A. F., & Hadders-Algra, M. (2010). Construct validity of the Infant Motor Profile: relation with prenatal, perinatal, and neonatal risk factors. *Developmental Medicine and Child Neurology*, *52*(9), 209–215. <https://doi.org/10.1111/j.1469-8749.2010.03667.x>
- Heinze, F., Hesels, K., Breitbach-Faller, N., Schmitz-Rode, T., & Disselhorst-Klug, C. (2010). Movement analysis by accelerometry of newborns and infants for the early detection of movement disorders due to infantile cerebral palsy. *Medical & Biological Engineering & Computing*, *48*(8), 765–772. <https://doi.org/10.1007/s11517-010-0624-z>
- Helenius, K., Sjörs, G., Shah, P. S., Modi, N., Reichman, B., Morisaki, N., ... Lehtonen, L. (2018). Survival in very preterm infants: An international comparison of 10 national neonatal networks. *Obstetrical and Gynecological Survey*, *73*(4), 187–189. <https://doi.org/10.1097/OGX.0000000000000553>
- Herber-Jonat, S., Streiftau, S., Knauss, E., Voigt, F., Flemmer, A. W., Hummler, H. D., ... Bode, H. (2014). Long-term outcome at age 7-10 years after extreme prematurity-a prospective, two centre cohort study of children born before 25 completed weeks of gestation (1999-2003). *Journal of Maternal-Fetal and Neonatal Medicine*, *27*(16), 1620–1626. <https://doi.org/10.3109/14767058.2013.871699>
- Herskind, A., Greisen, G., & Nielsen, J. B. (2014). Early identification and intervention in cerebral palsy. *Developmental Medicine & Child Neurology*, *57*(1), 29–36. <https://doi.org/10.1111/dmcn.12531>
- Hines, D., Modi, N., Lee, S. K., Isayama, T., Sjörs, G., Gagliardi, L., ... Shah, P. S. (2017). Scoping review shows wide variation in the definitions of bronchopulmonary dysplasia in preterm infants and calls for a consensus. *Acta Paediatrica, International Journal of Paediatrics*, *106*(3), 366–374. <https://doi.org/10.1111/apa.13672>
- Hintz, S., Kendrick, D., Stoll, B., Vohr, B., Fanaroff, A., Donovan, E., ... Higgins, R. (2005). Neurodevelopmental and Growth Outcomes of Extremely Low Birth Weight Infants After Necrotizing Enterocolitis. *Pediatrics*, *115*(3), 696–703. <https://doi.org/10.1542/peds.2004-0569>
- Hintz, S. R., Kendrick, D. E., Vohr, B. R., & Poole, W. K. (2005). Changes in neurodevelopmental outcomes at 18 to 22 months corrected age among infants of less than 25 weeks' gestational age born in 1993-1999. *Obstetrical and Gynecological Survey*, *60*(11), 714–715. <https://doi.org/10.1097/01.ogx.0000182902.54465.2e>
- Ihlen, E. A. F., Støen, R., Boswell, L., de Regnier, R.-A., Fjørtoft, T., Gaebler-Spira, D., ... Adde, L. (2019). Machine Learning of Infant Spontaneous Movements for the Early Prediction of Cerebral Palsy: A Multi-Site Cohort Study. *Journal of Clinical Medicine*, *9*(1), 5. <https://doi.org/10.3390/jcm9010005>
- Imamura, T., Ariga, H., Kaneko, M., Watanabe, M., Shibukawa, Y., Fukuda, Y., ...

- Fujiki, T. (2013). Neurodevelopmental outcomes of children with periventricular leukomalacia. *Pediatrics and Neonatology*, *54*(6), 367–372. <https://doi.org/10.1016/j.pedneo.2013.04.006>
- Inder, T. E. (2005). Abnormal Cerebral Structure Is Present at Term in Premature Infants. *Pediatrics*, *115*(2), 286–294. <https://doi.org/10.1542/peds.2004-0326>
- Ismail, F. Y., Fatemi, A., & Johnston, M. V. (2017). Cerebral plasticity: Windows of opportunity in the developing brain. *European Journal of Paediatric Neurology*, *21*(1), 23–48. <https://doi.org/10.1016/j.ejpn.2016.07.007>
- Jacob, J., Lehne, M., Mischker, A., Klinger, N., Zickermann, C., & Walker, J. (2017). Cost effects of preterm birth: a comparison of health care costs associated with early preterm, late preterm, and full-term birth in the first 3 years after birth. *European Journal of Health Economics*, *18*(8), 1041–1046. <https://doi.org/10.1007/s10198-016-0850-x>
- Jaeschke, R., Guyatt, G., & Sackett, D. (1994). Users' Guides to the Medical Literature A. Are the Results of the. *Jama*, *271*, 389–391. Retrieved from JAMA User's Guide to the Medical Literature series
- Jarjour, I. T. (2015). Neurodevelopmental outcome after extreme prematurity: A review of the literature. *Pediatric Neurology*, *52*(2), 143–152. <https://doi.org/10.1016/j.pediatrneurol.2014.10.027>
- Jeng, S.-F., Kuo-Inn Tsou Yau, Chen, L., & Hsiao, S. (2000). Alberta Infant Motor Scale: Reliability and Validity When Used on Preterm Infants in Taiwan. *Physical Therapy*, *80*(2). <https://doi.org/10.1093/ptj/80.2.168>
- Jeong, S. U., Kim, G. C., Jeong, H. J., Kim, D. K., Hong, Y. R., Kim, H. D., ... Sim, Y.-J. (2019). The Validity of the Bayley-III and DDST-II in Preterm Infants With Neurodevelopmental Impairment : A Pilot Study, *41*(5), 851–857.
- Jobe, A. H., & Bancalari, E. (2001). Bronchopulmonary Dysplasia. *Am J Respir Crit Care Med*, *163*, 1723–1729.
- Jobe, A. H., & Steinhorn, R. (2017). Can We Define Bronchopulmonary Dysplasia? *Journal of Pediatrics*, *188*, 19–23. <https://doi.org/10.1016/j.jpeds.2017.06.064>
- Johnson, S., Evans, T. A., Draper, E. S., Field, D. J., Manktelow, B. N., Marlow, N., ... Boyle, E. M. (2015). Neurodevelopmental outcomes following late and moderate prematurity: A population-based cohort study. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *100*(4), F301–F308. <https://doi.org/10.1136/archdischild-2014-307684>
- Kanemaru, N., Watanabe, H., Kihara, H., Nakano, H., Nakamura, T., Nakano, J., ... Konishi, Y. (2014). Jerky spontaneous movements at term age in preterm infants who later developed cerebral palsy. *Early Human Development*, *90*(8), 387–392. <https://doi.org/10.1016/j.earlhumdev.2014.05.004>
- Kanemaru, N., Watanabe, H., Kihara, H., Nakano, H., Takaya, R., Nakamura, T., ... Konishi, Y. (2013). Specific characteristics of spontaneous movements in preterm infants at term age are associated with developmental delays at age 3 years. *Developmental Medicine & Child Neurology*, n/a-n/a.

<https://doi.org/10.1111/dmcn.12156>

- Karch, D., Kang, K. S., Wochner, K., Philippi, H., Hadders-Algra, M., Pietz, J., & Dickhaus, H. (2012). Kinematic assessment of stereotypy in spontaneous movements in infants. *Gait and Posture*, *36*(2), 307–311. <https://doi.org/10.1016/j.gaitpost.2012.03.017>
- Karch, D., Kim, K., Wochner, K., Pietz, J., Dickhaus, H., & Philippi, H. (2008). Quantification of the segmental kinematics of spontaneous infant movements, *41*, 2860–2867. <https://doi.org/10.1016/j.jbiomech.2008.06.033>
- Kelly, C. E., Thompson, D. K., Cheong, J. L. Y., Chen, J., Olsen, J. E., Eeles, A. L., ... Spittle, A. J. (2019). Brain structure and neurological and behavioural functioning in infants born preterm. *Developmental Medicine and Child Neurology*, *61*(7), 820–831. <https://doi.org/10.1111/dmcn.14084>
- Keszei, A. P., Novak, M., & Streiner, D. L. (2010). Introduction to health measurement scales. *Journal of Psychosomatic Research*, *68*(4), 319–323. <https://doi.org/10.1016/j.jpsychores.2010.01.006>
- Khan, A. M., Morris, S. K., & Bhutta, Z. A. (2017). Neonatal and Perinatal Infections. *Pediatric Clinics of North America*, *64*(4), 785–798. <https://doi.org/10.1016/j.pcl.2017.03.008>
- Khwaja, O., & Volpe, J. (2008). Pathogenesis of cerebral white matter injury of prematurity. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *93*(2), F153-61. <https://doi.org/10.1136/adc.2006.108837.Pathogenesis>
- Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Diseases*, *38*(1), 27–36. [https://doi.org/10.1016/0021-9681\(85\)90005-0](https://doi.org/10.1016/0021-9681(85)90005-0)
- Koivisto, M., Ilkko, E., Vainionpää, L., Pääkö, E., Lanning, F., & Valkama, A. (2010). Magnetic resonance imaging at term and neuromotor outcome in preterm infants. *Acta Paediatrica*, *89*(3), 348–355. <https://doi.org/10.1111/j.1651-2227.2000.tb01337.x>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kostovic, Goran, Vulksic, & Judas. (2015). The Relevance of Human Fetal Subplate Zone for Developmental Neuropathology of Neuronal Migration Disorders and Cortical Dysplasia Sequential Development and Transient Cellular Compartments of the Human Fetal Cerebral Wall Structural and Functional Organi, *21*(3), 74–82. <https://doi.org/10.1111/cns.12333>
- Kottner, J., Audig, L., Brorson, S., Donner, A., & Gajewski, B. J. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed, *64*, 96–106. <https://doi.org/10.1016/j.jclinepi.2010.03.002>
- Kraemer, H. C., Periyakoil, V. S., & Noda, A. (2002). Kappa coefficients in medical research. *Statistics in Medicine*, *21*(14), 2109–2129. <https://doi.org/10.1002/sim.1180>

- Kraft, K. E., Verhage, S. E., den Heijer, A. E., & Bos, A. F. (2019). Functional outcome at school age of preterm-born children treated with low-dose dexamethasone in infancy. *Early Human Development*, 129(December 2018), 16–22. <https://doi.org/10.1016/j.earlhumdev.2018.12.016>
- Krallis, N., Theocharis, P., Nikaki, A., Drougia, A., Tzoufi, M., Andronikou, S., & Giapros, V. (2007). Incidence and risk factors for cerebral palsy in infants with perinatal problems: A 15-year review. *Early Human Development*, 83(8), 541–547. <https://doi.org/10.1016/j.earlhumdev.2006.10.004>
- Krzywinski, M., & Altman, N. (2015). Points of Significance: Multiple linear regression. *Nature Methods*. Nature Publishing Group. <https://doi.org/10.1038/nmeth.3665>
- Kugelman, A., & Colin, A. A. (2013). Late Preterm Infants: Near Term But Still in a Critical Developmental Time Period. *Pediatrics*, 132(4), 741–751. <https://doi.org/10.1542/peds.2013-1131>
- Kugelman, Amir, Bader, D., Lerner-Geva, L., Boyko, V., Levitzki, O., Riskin, A., & Reichman, B. (2012). Poor outcomes at discharge among extremely premature infants: A national population-based study. *Archives of Pediatrics and Adolescent Medicine*, 166(6), 543–550. <https://doi.org/10.1001/archpediatrics.2011.891>
- Kurjak, A., Tikvica, A., Stanojevic, M., Miskovic, B., Ahmed, B., Azumendi, G., & Renzo, G. C. Di. (2008). The assessment of fetal neurobehavior by three-dimensional and four-dimensional ultrasound. *Journal of Maternal-Fetal and Neonatal Medicine*, 21(10), 675–684. <https://doi.org/10.1080/14767050802212166>
- Kwong, A. K. L., Eeles, A. L., Olsen, J. E., Cheong, J. L. Y., Doyle, L. W., & Spittle, A. J. (2019). The Baby Moves smartphone app for General Movements Assessment: Engagement amongst extremely preterm and term-born infants in a state-wide geographical study. *Journal of Paediatrics and Child Health*, 55(5), 548–554. <https://doi.org/10.1111/jpc.14240>
- Kwong, A. K. L., Fitzgerald, T. L., Doyle, L. W., Cheong, J. L. Y., & Spittle, A. J. (2018). Predictive validity of spontaneous early infant movement for later cerebral palsy: a systematic review. *Developmental Medicine and Child Neurology*, 60(5), 480–489. <https://doi.org/10.1111/dmcn.13697>
- Laptook, A. R. (2013). Neurologic and Metabolic Issues in Moderately Preterm, Late Preterm, and Early Term Infants. *Clinics in Perinatology*, 40(4), 723–738. <https://doi.org/10.1016/j.clp.2013.07.005>
- Larroque, B., Ancel, P.-Y., Marret, S., Marchand, L., André, M., Arnaud, C., ... Kaminski, M. (2008). Neurodevelopmental disabilities an special care of 5-years-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. *The Lancet*, 371(2), 813–820.
- Larroque, B., Bréart, G., Kaminski, M., Dehan, M., André, M., Burguet, A., ... Truffert, P. (2004). Survival of very preterm infants: Epipage, a population based cohort study. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 89(2), 136

139F – 144. <https://doi.org/10.1136/adc.2002.020396>

- Lawn, J. E., Stoll, B. J., Vergnano, S., Zaidi, A., Seale, A. C., Syed, S., ... Cousens, S. N. (2013). Neonatal severe bacterial infection impairment estimates in South Asia, sub-Saharan Africa and Latin America for 2010. *Pediatric Research*, 74(S1), 73–85. <https://doi.org/10.1038/pr.2013.207>
- Lee, I., Neil, J. J., Huettner, P. C., Smyser, C. D., Rogers, C. E., Shimony, J. S., ... E., T. (2014). The impact of prenatal and neonatal infection on neurodevelopmental outcomes in very preterm infants, 34(10), 741–747. <https://doi.org/10.1038/jp.2014.79>.The
- Lijmer, J. G. (1999). Empirical Evidence of Design-Related Bias. *Jama*, 282(11), 1061–1066.
- Lin, P. W., & Stoll, B. J. (2006). Necrotising enterocolitis. *Lancet*, 368(9543), 1271–1283. [https://doi.org/10.1016/S0140-6736\(06\)69525-1](https://doi.org/10.1016/S0140-6736(06)69525-1)
- Lipkin, P. H. (2005). Towards creation of a unified view of the neurodevelopment of the infant. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 103–106. <https://doi.org/10.1002/mrdd.20057>
- Liu, L., Johnson, H. L., Cousens, S., Perin, J., Scott, S., Lawn, J. E., ... Black, R. E. (2012). Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000. *The Lancet*, 379(9832), 2151–2161. [https://doi.org/10.1016/S0140-6736\(12\)60560-1](https://doi.org/10.1016/S0140-6736(12)60560-1)
- Lockman, J. J., & Thelen, E. (1993). Developmental Biodynamics: Brain, Body, Behavior Connections. *Child Development*, 64(4), 953–959. <https://doi.org/10.1111/j.1467-8624.1993.tb04181.x>
- Lohr, K. N., Aaronson, N. K., Alonso, J., Burnam, M. A., Patrick, D. L., Perrin, E. B., & Roberts, J. S. (1996). Evaluating quality-of-life and health status instruments: Development of scientific review criteria. *Clinical Therapeutics*, 18(5), 979–984. [https://doi.org/10.1016/S0149-2918\(96\)80054-3](https://doi.org/10.1016/S0149-2918(96)80054-3)
- Long, T., & Tieman, B. (1998). Review of Two Recently Published Measurement tools : The AIMS ans the TIME.
- Lüchinger, A., Hadders-Algra, M., van Kan, C., & de Vries, J. (2008). Fetal Onset of General Movements. *PEDIATRIC RESEARCH*, 63(2), 191–195.
- Mackay-Lyons, M. (2002). Central Pattern Generation of Locomotion : A Review of the Evidence. *Physical Therapy*, 82(1), 69–83.
- Maitre, N. L., Marshall, D. D., Goldstein, R. F., Slaughter, J. C., & Price, W. A. (2011). Necrotizing enterocolitis in infants with periventricular hemorrhagic infarction: Associations and outcomes. *Neonatology*, 99(2), 97–103. <https://doi.org/10.1159/000313960>
- Maitre N.L., Nelin M.A., Noritz G.N., Chorna O.C., Williams J.W., Carey H.C., G. A. . (2017). Routine use of the general movements assessment in a US neonatal intensive care unit to improve targeted neuroimaging and follow-up of infants at high-risk for movement disorders. *Developmental Medicine & Child Neurology*,

59, 46–87. <https://doi.org/10.1111/dmcn.13456>

- Majnemer, A., & Mazer, B. (1998). Neurologic evaluation of the newborn infant: Definition and psychometric properties. *Developmental Medicine and Child Neurology*, 40, 708–715. <https://doi.org/10.1111/j.1469-8749.1998.tb12332.x>
- Majnemer, A., Riley, P., Shevell, M., Birnbaum, R., Greenstone, H., & Coates, a L. (2000). Severe bronchopulmonary dysplasia increases risk for later neurological and motor sequelae in preterm survivors. *Developmental Medicine and Child Neurology*, 42(1), 53–60. <https://doi.org/10.1017/S001216220000013X>
- Mandrekar, J. N. (2011). Measures of interrater agreement. *Journal of Thoracic Oncology*, 6(1), 6–7. <https://doi.org/10.1097/JTO.0b013e318200f983>
- Mansson, J., & Stjernqvist, K. (2014). Children born extremely preterm show significant lower cognitive , language and motor function levels compared with children born at term , as measured by the Bayley-III at 2 . 5 years, 504–511. <https://doi.org/10.1111/apa.12585>
- Marcroft, C., Khan, A., Embleton, N. D., Trenell, M., & Plötz, T. (2015). Movement Recognition Technology as a Method of Assessing Spontaneous General Movements in High Risk Infants. *Frontiers in Neurology*, 5(January). <https://doi.org/10.3389/fneur.2014.00284>
- Marill, K. A. (2004). Advanced Statistics : Multiple Linear Regression. *Acad Emerg Med*, 11(1), 94–102. [https://doi.org/10.1197/S1069-6563\(03\)00601-8](https://doi.org/10.1197/S1069-6563(03)00601-8)
- Marschik, P. B., Pokorny, F. B., Peharz, R., Zhang, D., O’Muircheartaigh, J., Roeyers, H., ... Kaufmann, W. E. (2017). A Novel Way to Measure and Predict Development: A Heuristic Approach to Facilitate the Early Detection of Neurodevelopmental Disorders. *Current Neurology and Neuroscience Reports*, 17(5). <https://doi.org/10.1007/s11910-017-0748-8>
- McCormick, D. (1999). Spontaneous Activity: Signal or Noise?, (July).
- McCormick, M. (1985). The contribution of low birth weight to infant mortality and childhood morbidity, 312.
- McCrea, H. J., Ment, L. R., & McCrea Heather and Ment Laura. (2010). The Diagnosis, Management and Postnatal Prevention of Intraventricular Hemorrhage in the Preterm Neonate. *Prevention*, 35(4), 777–792, vii. <https://doi.org/10.1016/j.clp.2008.07.014>.The
- McHugh, M. L. (2012). Lessons in biostatistics interrater reliability : the kappa statistic. *Biochemica Medica*, 22(3), 276–282. Retrieved from <https://hrcak.srce.hr/89395>
- McKenzie, D. P., Vida, S., Mackinnon, A. J., Onghena, P., & Clarke, D. M. (1997). Accurate confidence intervals for measures of test performance. *Psychiatry Research*, 69(2–3), 207–209. [https://doi.org/10.1016/S0165-1781\(96\)02952-6](https://doi.org/10.1016/S0165-1781(96)02952-6)
- Meehan, C. F., Grondahl, L., Nielsen, J. B., & Hultborn, H. (2012). Fictive locomotion in the adult decerebrate and spinal mouse in vivo. *Journal of Physiology*, 590(2), 289–300. <https://doi.org/10.1113/jphysiol.2011.214643>

- Meinecke, L., Breitbach-Faller, N., Bartz, C., Damen, R., Rau, G., & Disselhorst-Klug, C. (2006). Movement analysis in the early detection of newborns at risk for developing spasticity due to infantile cerebral palsy. *Human Movement Science, 25*(2), 125–144. <https://doi.org/10.1016/j.humov.2005.09.012>
- Ment, L. R., Bada, H. S., Barnes, P., Grant, P. E., Hirtz, D., Papile, L. A., ... Slovis, T. L. (2012). Practice parameter: Neuroimaging of the neonate: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology, 59*(10), 1663–1664. <https://doi.org/10.1212/wnl.59.10.1663>
- Merhar, S. L., Tabangin, M. E., Meinzen-Derr, J., & Schibler, K. R. (2012). Grade and Laterality of Intraventricular Hemorrhage to Predict 18–22 Month Neurodevelopmental Outcomes in Extremely Low Birth Weight Infants, *101*(4), 414–418. <https://doi.org/10.1111/j.1651-2227.2011.02584.x>.Grade
- Meurer, W. J., & Tolles, J. (1986). Logistic Regression Diagnostics Understanding HowWell a Model Predicts Outcomes. *Journal of the American Statistical Association, 81*(394), 461. <https://doi.org/10.2307/2289236>
- Mewes, A. U. J., Huppi, P. S., Als, H., Rybicki, F. J., Inder, T. E., McAnulty, G. B., ... Warfield, S. K. (2006). Regional Brain Development in Serial Magnetic Resonance Imaging of Low-Risk Preterm Infants. *Pediatrics, 118*(1), 23–33. <https://doi.org/10.1542/peds.2005-2675>
- Milligan, D. (2010). Outcomes of children born very preterm in Europe. *Archives of Disease in Childhood: Fetal and Neonatal Edition, 95*(4). <https://doi.org/10.1136/adc.2008.143685>
- Minassian, K., Hofstoetter, U. S., Dzeladini, F., Guertin, P. A., & Ijspeert, A. (2017). The Human Central Pattern Generator for Locomotion: Does It Exist and Contribute to Walking? *Neuroscientist, 23*(6), 649–663. <https://doi.org/10.1177/1073858417699790>
- Mirmiran, M. (2004). Neonatal Brain Magnetic Resonance Imaging Before Discharge Is Better Than Serial Cranial Ultrasound in Predicting Cerebral Palsy in Very Low Birth Weight Preterm Infants. *Pediatrics, 114*(4), 992–998. <https://doi.org/10.1542/peds.2003-0772-1>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. (2012). The COSMIN checklist manual. ... : *VU University Medical ...*, 56. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... De Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life*

Research, 19(4), 539–549. <https://doi.org/10.1007/s11136-010-9606-8>

- Moore, A., Zhou, W., Jakovcevski, I., Zecevic, N., & Antic, S. (2011). Spontaneous electrical activity in the human fetal cortex in vitro. *The Journal of Neuroscience*, 31(7), 2391–2398. <https://doi.org/10.1523/JNEUROSCI.3886-10.2011>. Spontaneous
- Moore, T., Hennessy, E. M., Myles, J., Johnson, S. J., Draper, E. S., Costeloe, K. L., & Marlow, N. (2013). Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: The EPICURE studies. *Obstetrical and Gynecological Survey*, 68(4), 274–275. <https://doi.org/10.1097/01.OGX.0000429293.80760.f2>
- Morgan, J. C., & Boyle, E. M. (2018). The late preterm infants. *Paediatrica Croatica, Supplement*, 62, 32–38. <https://doi.org/10.1016/j.paed.2017.10.003>
- Mukerji, A., Shah, V., & Shah, P. S. (2015). Periventricular/Intraventricular Hemorrhage and Neurodevelopmental Outcomes: A Meta-analysis. *Pediatrics*, 136(6), 1132–1143. <https://doi.org/10.1542/peds.2015-0944>
- Murthy, V., & Kennea, N. L. (2007). Antenatal infection/inflammation and fetal tissue injury. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, 21(3), 479–489. <https://doi.org/10.1016/j.bpobgyn.2007.01.010>
- Nagara, S., Kouwaki, M., Togawa, T., Sugiura, T., Okada, M., & Koyama, N. (2018). Neurodevelopmental outcomes at 3 years old for infants with birth weights under 500 g. *Pediatrics and Neonatology*, 59(3), 274–280. <https://doi.org/10.1016/j.pedneo.2017.09.005>
- Nakajima, Y., Einspieler, C., Marschik, P. B., Bos, A. F., & Prechtl, H. F. R. (2006). Does a detailed assessment of poor repertoire general movements help to identify those infants who will develop normally? *Early Human Development*, 82(1), 53–59. <https://doi.org/10.1016/j.earlhumdev.2005.07.010>
- Natarajan, G., & Shankaran, S. (2016). Short- and long-term outcomes of moderate and late preterm infants. *American Journal of Perinatology*, 33(03), 305–317. <https://doi.org/10.1055/s-0035-1571150>
- Nelson, K. B., & Ellenberg, J. H. (1981). Apgar Scores as Predictors of Chronic Neurologic Disability. *Pediatrics*, 68(1), 36 LP – 44. Retrieved from <http://pediatrics.aappublications.org/content/68/1/36.abstract>
- Neu, J. (2014). Necrotizing enterocolitis. *World Rev Nutr Diet*, 110, 253–263. <https://doi.org/10.4324/9781315113968>
- Nijhuis, J. G., Swaab, Visser, Heuser, Finch, Brosens, & Eskes. (2003). Fetal behavior. *Neurobiology of Aging*, 24(SUPPL. 1), 3–8. [https://doi.org/10.1016/S0197-4580\(03\)00054-X](https://doi.org/10.1016/S0197-4580(03)00054-X)
- Noble, Y., & Boyd, R. (2012). Neonatal assessments for the preterm infant up to 4 months corrected age: a systematic review. *Developmental Medicine & Child Neurology*, 54(2), 129–139. <https://doi.org/10.1111/j.1469-8749.2010.03903.x>
- Nongena, P., Ederies, A., Azzopardi, D. V., & Edwards, A. D. (2010). Confidence in

- the prediction of neurodevelopmental outcome by cranial ultrasound and MRI in preterm infants. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 95(6), 388–391. <https://doi.org/10.1136/adc.2009.168997>
- Novak, C. M., Ozen, M., & Burd, I. (2018). Perinatal Brain Injury: Mechanisms, Prevention, and Outcomes. *Clinics in Perinatology*, 45(2), 357–375. <https://doi.org/10.1016/j.clp.2018.01.015>
- Novak, I., Morgan, C., Adde, L., Blackman, J., Boyd, R. N., Brunstrom-Hernandez, J., ... Badawi, N. (2017). Early, accurate diagnosis and early intervention in cerebral palsy: Advances in diagnosis and treatment. *JAMA Pediatrics*, 171(9), 897–907. <https://doi.org/10.1001/jamapediatrics.2017.1689>
- O'Shea, T. M., Klinepeter, K. L., Meis, P. J., & Dillard, R. G. (1998). Intrauterine infection and the risk of cerebral palsy in very low-birthweight infants. *Paediatric and Perinatal Epidemiology*, 12(1), 72–83. <https://doi.org/10.1111/j.1365-3016.1998.00081.x>
- O'Shea, T. M., Kuban, K. C. K., Allred, E. N., Paneth, N., Pagano, M., Dammann, O., ... Levi, A. (2010). Neonatal Cranial Ultrasound Lesions and Developmental Delays at 2 Years of Age Among Extremely Low Gestational Age Children, 122(3). <https://doi.org/10.1542/peds.2008-0594.Neonatal>
- Ohgi, S., Morita, S., Loo, K. K., & Mizuike, C. (2008). Time series analysis of spontaneous upper-extremity movements of premature infants with brain injuries. *Physical Therapy*, 88(9), 1022–1033. <https://doi.org/10.2522/ptj.20070171>
- Okado, N. (1980). Development of the human cervical spinal cord with reference to synapse formation in the motor nucleus. *Journal of Comparative Neurology*, 191(3), 495–513. <https://doi.org/10.1002/cne.901910311>
- Olsen, J.E., Brown, N. C., Eeles, A. L., Einspieler, C., Lee, K. J., Thompson, D. K., ... Spittle, A. J. (2016). Early general movements and brain magnetic resonance imaging at term-equivalent age in infants born < 30 weeks' gestation. *Early Human Development*, 101, 63–68. <https://doi.org/10.1016/j.earlhumdev.2016.06.009>
- Olsen, Joy E, Brown, N. C., Eeles, A. L., Lee, K. J., Anderson, P. J., Cheong, J. L. Y., ... Spittle, A. J. (2015). Trajectories of general movements from birth to term-equivalent age in infants born b 30 weeks ' gestation. *Early Hum Dev*, 91, 683–688.
- Orchinik, L. J., Taylor, H. G., Espy, K. A., Minich, N., Klein, N., Sheffield, T., & Hack, M. (2008). Cognitive Outcomes for Extremely Preterm/Extremely Low Birth Weight Children in Kindergarten. *Prog Retin Eye Res*, 26(1), 38–56. <https://doi.org/10.1017/S135561771100107X.Cognitive>
- Palmer, F. B. (2004). Strategies for the early diagnosis of cerebral palsy. *The Journal of Pediatrics*, 145(2), S8–S11. <https://doi.org/10.1016/j.jpeds.2004.05.016>
- Papile, M.D. (1978). Incidence and evolution of subependymal and intraventricular hemorrhage: A study of infants with birth weights less than 1 , 500 gm. *The*

Journal of Pediatrics, 92(4), 529–534.

- Parikh, N. A. (2016). Advanced neuroimaging and its role in predicting neurodevelopmental outcomes in very preterm infants. *Semin Perinatol*, 20(2), 163–178. <https://doi.org/10.1007/s10741-014-9462-7>. Natural
- Park, H. W., Yoon, H. K., Han, S. B., Lee, B. S., Sung, I. Y., Kim, K. S., & Kim, E. A. R. (2014). Brain MRI measurements at a term-equivalent age and their relationship to neurodevelopmental outcomes. *American Journal of Neuroradiology*, 35(3), 599–603. <https://doi.org/10.3174/ajnr.A3720>
- Parniere, C. (2017). Incertitude diagnostique : quels ressentis pour les patients ?
- Paro-Panjan, D., Sustersic, B., & Neubauer, D. (2005). Comparison of two methods of neurologic assessment in infants. *Pediatric Neurology*, 33(5), 317–324. <https://doi.org/10.1016/j.pediatrneurol.2005.05.008>
- Pascal, A., Govaert, P., & Broeck, C. V. A. N. D. E. N. (2018). Neurodevelopmental outcome in very preterm and very-low- birthweight infants born over the past decade : a meta-analytic review, 1–14. <https://doi.org/10.1111/dmcn.13675>
- Pascal, A., Govaert, P., Oostra, A., Naulaers, G., Ortibus, E., & Van den Broeck, C. (2018). Neurodevelopmental outcome in very preterm and very-low-birthweight infants born over the past decade: a meta-analytic review. *Developmental Medicine and Child Neurology*, 60(4), 342–355. <https://doi.org/10.1111/dmcn.13675>
- Patel, R. M. (2016). Short- and Long-Term Outcomes for Extremely Preterm Infants. *American Journal of Perinatology*, 33(3), 318–328. <https://doi.org/10.1055/s-0035-1571202>
- Patra, K., Wilson-Costello, D., Taylor, G., Mercuri-Minich, N., & Hack, M. (2006). GRADES I-II INTRAVENTRICULAR HEMORRHAGE IN EXTREMELY LOW BIRTH WEIGHT INFANTS: EFFECTS ON NEURODEVELOPMENT, 169–173.
- Peyton, C., Yang, E., Kocherginsky, M., Adde, L., Fjørtoft, T., Støen, R., ... Msall, M. E. (2016). Relationship between white matter pathology and performance on the General Movement Assessment and the Test of Infant Motor Performance in very preterm infants. *Early Human Development*, 95, 23–27. <https://doi.org/10.1016/j.earlhumdev.2016.01.017>
- Peyton, Colleen, & Einspieler, C. (2018). General Movements: A Behavioral Biomarker of Later Motor and Cognitive Dysfunction in NICU Graduates. *Pediatric Annals*, 47(4), e159–e164. <https://doi.org/10.3928/19382359-20180325-01>
- Pfahl, S., Hoehn, T., Lohmeier, K., Richter-Werkle, R., Babor, F., Schramm, D., & Sabir, H. (2018). Long-term neurodevelopmental outcome following low grade intraventricular hemorrhage in premature infants. *Early Human Development*, 117(October 2017), 62–67. <https://doi.org/10.1016/j.earlhumdev.2017.12.013>
- Philippi, H., Karch, D., Kang, K.-S., Wochner, K., Pietz, J., Dickhaus, H., & Hadders-Algra, M. (2014). Computer-based analysis of general movements reveals stereotypies predicting cerebral palsy. *Developmental Medicine & Child*

- Neurology*, 56(10), 960–967. <https://doi.org/10.1111/dmcn.12477>
- Pierrat, V., Marchand-Martin, L., Arnaud, C., Kaminski, M., Resche-Rigon, M., Lebeaux, C., ... Ancel, P.-Y. (2017). Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in France in 2011: EPIPAGE-2 cohort study. *Bmj*, j3448. <https://doi.org/10.1136/bmj.j3448>
- Pin, T. W., De Valle, K., Eldridge, B., & Galea, M. P. (2010). Clinimetric properties of the alberta infant motor scale in infants born preterm. *Pediatric Physical Therapy*, 22(3), 278–286. <https://doi.org/10.1097/PEP.0b013e3181e94481>
- Polit, D. F., & Beck, C. T. (2006). The Content Validity Index : Are You Sure You Know What ' s Being Reported ? Critique and Recommendations, 489–497. <https://doi.org/10.1002/nur>
- Potharst, E. S., Van Wassenaer-Leemhuis, A. G., Houtzager, B. A., Livesey, D., Kok, J. H., Last, B. F., & Oosterlaan, J. (2013). Perinatal risk factors for neurocognitive impairments in preschool children born very preterm. *Developmental Medicine and Child Neurology*, 55(2), 178–184. <https://doi.org/10.1111/dmcn.12018>
- Prechtl, H. (1990). Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Human Development*, 23(3), 151–158. [https://doi.org/10.1016/0378-3782\(90\)90011-7](https://doi.org/10.1016/0378-3782(90)90011-7)
- Prechtl, H., Einspieler, C., Cioni, G., Bos, a F., Ferrari, F., & Sontheimer, D. (1997). An early marker for neurological deficits after perinatal brain lesions. *Lancet*, 349(9062), 1361–1363. [https://doi.org/10.1016/S0140-6736\(96\)10182-3](https://doi.org/10.1016/S0140-6736(96)10182-3)
- Prechtl, H. F. (2001). General movement assessment as a method of developmental neurology: new paradigms and their consequences. The 1999 Ronnie MacKeith lecture. *Developmental Medicine and Child Neurology*, 43(12), 836–842. <https://doi.org/10.1111/j.1469-8749.2001.tb00173.x>
- Prechtl, H. F. R. (1974). The behavioural states of the newborn infant (a review). *Brain Research*, 76(September 1973). [https://doi.org/10.1016/0006-8993\(74\)90454-5](https://doi.org/10.1016/0006-8993(74)90454-5)
- Prechtl, H. F. R. (1997). State of the art of a new functional assessment of the young nervous system. An early predictor of cerebral palsy. *Early Human Development*, 50(1), 1–11. [https://doi.org/10.1016/S0378-3782\(97\)00088-1](https://doi.org/10.1016/S0378-3782(97)00088-1)
- Rahmanpour, P. (2009). Features for movement based prediction of Cerebral Palsy. 67, (June). Retrieved from <https://brage.bibsys.no/xmlui/handle/11250/259906>
- Ranke, M. B., Krägeloh-Mann, I., & Vollmer, B. (2015). Growth, head growth, and neurocognitive outcome in children born very preterm: Methodological aspects and selected results. *Developmental Medicine and Child Neurology*, 57(1), 23–28. <https://doi.org/10.1111/dmcn.12582>
- Rau, M. P., Lin, Y. C., Ko, Y. C., & Hsue, B. J. (2014). Inter-rater and intra-rater reliability of the qualitative assessment of general movements on the infants with typical development and high-risk factors for developmental delay in Taiwan. *Annals of Physical and Rehabilitation Medicine*, 57, e345–e346.

<https://doi.org/10.1016/j.rehab.2014.03.1266>

- Rees, C. M., Pierro, A., & Eaton, S. (2007). Neurodevelopmental outcomes of neonates with medically and surgically treated necrotizing enterocolitis. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 92(3), 193–198. <https://doi.org/10.1136/adc.2006.099929>
- Ricci, E., Einspieler, C., & Craig, A. K. (2017). Feasibility of Using the General Movements Assessment of Infants in the United States. *Physical & Occupational Therapy in Pediatrics*, 2638(November), 1–11. <https://doi.org/10.1080/01942638.2017.1395380>
- Rich, B., & Stephen, D. (2017). Necrotizing enterocolitis. *Newborn Surgery, Fourth Edition*, 653–659. <https://doi.org/10.4324/9781315113968>
- Roach, K. E. (2006). Measurement of Health Outcomes : Reliability , Validity and Responsiveness. *J Prosthet Orthot*, 18(1S), 8–12.
- Roescher, M., Timmer, A., Hitzert, M. M., de Vries, N. K. S., Verhagen, E. a, Erwich, J. J. H. M., & Bos, a F. (2014). Placental pathology and neurological morbidity in preterm infants during the first two weeks after birth. *Early Human Development*, 90(1), 21–25. <https://doi.org/10.1016/j.earlhumdev.2013.11.004>
- Rogers, E. E., & Hintz, S. R. (2016). Early neurodevelopmental outcomes of extremely preterm infants. *Seminars in Perinatology*, 40(8), 497–509. <https://doi.org/10.1053/j.semperi.2016.09.002>
- Romeo, D M, Luciano, R., Corsello, M., Ricci, D., Brogna, C., Zuppa, A., ... Mercuri, E. (2013). Neonatal neurological examination of late preterm babies. *Early Human Development*, 89(8), 537–545. Retrieved from <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medl&AN=23380499%5Cnhttp://202.120.227.56:3210/sfxlcl3?sid=OVID:m edline&id=pmid:23380499&id=doi:10.1016/j.earlhumdev.2013.01.002&issn=0378-3782&isbn=&volume=89&issue=8&spage=537&pages=53>
- Romeo, Domenico M., Ricci, D., Brogna, C., Cilauro, S., Lombardo, M. E., Romeo, M. G., & Mercuri, E. (2011). Neurological examination of late-preterm infants at term age. *European Journal of Paediatric Neurology*, 15(4), 353–360. <https://doi.org/10.1016/j.ejpn.2011.03.009>
- Rose, R. U., & Westcott, S. L. (2005). Responsiveness of the Test of Infant Motor Performance (TIMP) in infants born preterm. *Pediatric Physical Therapy*, 17(3), 219–224. <https://doi.org/10.1097/01.PEP.0000176575.63915.67>
- Rosenbaum, P. (2006). Variation and “abnormality”: recognizing the differences. *The Journal of Pediatrics*, 2005–2006.
- Sarkar, S., Shankaran, S., Barks, J., Do, B., Laptook, A., Abhik, D., ... Shriver, E. K. (2018). Outcome of Preterm Infants with Transient Cystic Periventricular Leukomalacia on Serial Cranial Imaging Up to Term Equivalent Age. *J Pediatr*, 263(2), 219–227. <https://doi.org/10.1097/SLA.0000000000001177.Complications>
- Schneider, A., Hommel, G., & Blettner, M. (2010). Lineare regressionsanalyse - Teil

- 14 der serie zur bewertung wissenschaftlicher publikationen. *Deutsches Arzteblatt*, 107(44), 776–782. <https://doi.org/10.3238/arztebl.2010.0776>
- Schulzke, S. M., Deshpande, G. C., & Patole, S. K. (2007). Neurodevelopmental Outcomes of Very Low-Birth-Weight Infants With Necrotizing Enterocolitis. *Archives of Pediatrics & Adolescent Medicine*, 161(6), 583. <https://doi.org/10.1001/archpedi.161.6.583>
- Seme-Ciglenecki, Polona. (2007). Predictive values of cranial ultrasound and assessment of general movements for neurological development of preterm infants in the Maribor region of Slovenia. *Wiener Klinische Wochenschrift*, 119(15–16), 490–496. <https://doi.org/10.1007/s00508-007-0839-7>
- Seme-Ciglenecki, Polonca. (2003). Predictive value of assessment of general movements for neurological development of high-risk preterm infants: comparative study. *Croatian Medical Journal*, 44(6), 721–727. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14652886
- Serenius, F., Ewald, U., Farooqi, A., Fellman, V., Hafström, M., Hellgren, K., ... Källén, K. (2016). Neurodevelopmental Outcomes Among Extremely Preterm Infants 6.5 Years After Active Perinatal Care in Sweden. *JAMA Pediatrics*, 170(10), 954. <https://doi.org/10.1001/jamapediatrics.2016.1210>
- Shah, P., Kaciroti, N., Richards, B., Oh, W., & Lumeng, J. C. (2016). Developmental Outcomes of Late Preterm Infants From Infancy to Kindergarten, 138(2). <https://doi.org/10.1542/peds.2015-3496>
- Shah, T. A., Meinzen-Derr, J., Gratton, T., Steichen, J., Donovan, E. F., Yolton, K., ... Schibler, K. R. (2012). Hospital and neurodevelopmental outcomes of extremely low-birth-weight infants with necrotizing enterocolitis and spontaneous intestinal perforation. *Journal of Perinatology*, 32(7), 552–558. <https://doi.org/10.1038/jp.2011.176>
- Shane, A. L., Sánchez, P. J., & Stoll, B. J. (2017). Neonatal sepsis. *Newborn Surgery, Fourth Edition*, 6736(17), 217–232. <https://doi.org/10.4324/9781315113968>
- Sherlock, L., & Abman, S. H. (2018). Bronchopulmonary Dysplasia. *Kendig's Disorders of the Respiratory Tract in Children*, 367-381.e5. <https://doi.org/10.1016/B978-0-323-44887-1.00020-1>
- Short, E. J., Klein, N. K., Lewis, B. A., Fulton, S., Eisengart, S., Kercksmar, C., ... Singer, L. T. (2003). Cognitive and Academic Consequences of Bronchopulmonary Dysplasia and Very Low Birth Weight: 8-Year-Old Outcomes, 112(5), 359–366. <https://doi.org/10.1542/peds.112.5.e359>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations : Uses in Assessing Rater Reliability, 86(2), 420–428.
- Shulhan, J., Dicken, B., Hartling, L., & Larsen, B. M. (2017). Current Knowledge of Necrotizing Enterocolitis in Preterm Infants and the Impact of Different Types of Enteral Nutrition Products. *Advances in Nutrition: An International Review*

Journal, 8(1), 80–91. <https://doi.org/10.3945/an.116.013193>

- Simon, S. D. (2001). Andrology Lab Corner Understanding the Odds Ratio and the Relative Risk. *Journal of Andrology*, 22(4). <https://doi.org/10.1111/j.1939-4640.2001.tb02212.x>
- Simpson, S. J., Hall, G. L., & Wilson, A. C. (2015). Lung function following very preterm birth in the era of “new” bronchopulmonary dysplasia. *Respirology*, 20(4), 535–540. <https://doi.org/10.1111/resp.12503>
- Singer, L., Yamashita, T., Lilien, L., Collin, M., & Jill, B. (1997). A Longitudinal Study of Developmental Outcome of Infants With Bronchopulmonary Dysplasia and Very Low Birth Weight, 100(6), 987–993. <https://doi.org/10.3899/jrheum.121180.Response>
- Skidmore, M. D., Rivers, A., & Hack, M. (1990). Increased Risk of Cerebral Palsy Among Very Low-Birthweight Infants With Chronic Lung Disease. *Developmental Medicine & Child Neurology*, 32(4), 325–332. <https://doi.org/10.1111/j.1469-8749.1990.tb16944.x>
- Snider, L. M., Majnemer, A., Mazer, B., Campbell, S., & Bos, A. F. (2008). A comparison of the general movements assessment with traditional approaches to newborn and infant assessment: Concurrent validity. *Early Human Development*, 84(5), 297–303. <https://doi.org/10.1016/j.earlhumdev.2007.07.004>
- Snider, L., Majnemer, A., Mazer, B., Campbell, S., & Bos, A. F. (2009). Prediction of motor and functional outcomes in infants born preterm assessed at term. *Pediatric Physical Therapy: The Official Publication of the Section on Pediatrics of the American Physical Therapy Association*, 21(1), 2–11. <https://doi.org/10.1097/PEP.0b013e3181957bdc>
- Song, Y. H., Chang, H. J., Shin, Y. B., Park, Y. S., Park, Y. H., & Cho, E. S. (2018). The validity of two neuromotor assessments for predicting motor performance at 12 months in preterm infants. *Annals of Rehabilitation Medicine*, 42(2), 296–304. <https://doi.org/10.5535/arm.2018.42.2.296>
- Sonntag, J., Grimmer, I., Scholz, T., Metze, B., Wit, J., & Obladen, M. (2010). Growth and neurodevelopmental outcome of very low birthweight infants with necrotizing enterocolitis. *Acta Paediatrica*, 89(5), 528–532. <https://doi.org/10.1111/j.1651-2227.2000.tb00332.x>
- Soraisham, A. S., Amin, H. J., Al-Hindi, M. Y., Singhal, N., & Sauve, R. S. (2006). Does necrotising enterocolitis impact the neurodevelopmental and growth outcomes in preterm infants with birthweight ≤ 1250 g? *Journal of Paediatrics and Child Health*, 42(9), 499–504. <https://doi.org/10.1111/j.1440-1754.2006.00910.x>
- Spittle, A., Doyle, L., & Boyd, R. (2008). A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Developmental Medicine & Child Neurology*, 50(4), 254–266. <https://doi.org/10.1111/j.1469-8749.2008.02025.x>

- Spittle, A. J., Olsen, J., Kwong, A., Doyle, L. W., Marschik, P. B., Einspieler, C., & Cheong, J. L. Y. (2016). The Baby Moves prospective cohort study protocol: Using a smartphone application with the General Movements Assessment to predict neurodevelopmental outcomes at age 2 years for extremely preterm or extremely low birthweight infants. *BMJ Open*, *6*(10), 1–8. <https://doi.org/10.1136/bmjopen-2016-013446>
- Spittle, A J, Brown, N. C., Doyle, L. W., Boyd, R. N., Hunt, R. W., Bear, M., & Inder, T. E. (2008). Quality of general movements is related to white matter pathology in very preterm infants. *Pediatrics*, *121*(5), e1184-9. <https://doi.org/10.1542/peds.2007-1924>
- Spittle, Alicia J., Walsh, J., Olsen, J. E., McInnes, E., Eeles, A. L., Brown, N. C., ... Cheong, J. L. Y. (2016). Neurobehaviour and neurological development in the first month after birth for infants born between 32-42 weeks' gestation. *Early Human Development*, *96*, 7–14. <https://doi.org/10.1016/j.earlhumdev.2016.02.006>
- Spittle, Alicia J, Boyd, R. N., Inder, T. E., & Doyle, L. W. (2009). Predicting motor development in very preterm infants at 12 months' corrected age: the role of qualitative magnetic resonance imaging and general movements assessments. *Pediatrics*, *123*(2), 512–517. <https://doi.org/10.1542/peds.2008-0590>
- Spittle, Alicia, Walsh, J., Potter, C., Mcinnes, E., Olsen, J., Lee, K., ... Cheong, J. (2017). Neurobehaviour at term-equivalent age and neurodevelopmental outcomes at 2 years in infants born moderate-to-late preterm. *Developmental Medicine and Child Neurology*, *59*(2), 207–215. <https://doi.org/10.1111/dmcn.13297>
- Stecina, K. (2017). Midbrain stimulation-evoked lumbar spinal activity in the adult decerebrate mouse. *Journal of Neuroscience Methods*, *288*, 99–105. <https://doi.org/10.1016/j.jneumeth.2017.06.015>
- Stephen W. Smoliar. (1989). Book Reviews: Gerald M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection* (Basic Books; New York, 1987); xxii + 371 pages. *ARTIFICIAL INTELLIGENCE*, *39*, 121–136. Retrieved from https://ac.els-cdn.com/0004370289900040/1-s2.0-0004370289900040-main.pdf?_tid=eff4a1dc-1dc0-44d2-991f-33d7adf55171&acdnat=1526354289_656e829e087a705603e890efa7059f3a
- Stöckl, D., Rodríguez, D., & Thienpont, L. (2004). Interpreting Method Comparison Studies by Use of the Bland–Altman Plot: Reflecting the Importance of Sample Size by Incorporating Confidence Limits and Predefined Error Limits in the Graphic. *Clinical Chemistry*, *50*(11), 2215–2216. <https://doi.org/10.1373/clinchem.2004.040162>
- Stoelhorst, M., Rijken, M., Martens, S. E., Brand, R., Ouden, A. L. den, Wit, J.-M., & Veen, S. (2005). Changes in Neonatology: Comparison of Two Cohorts of Very Preterm Infants (Gestational Age <32 Weeks): The Project on Preterm and Small for Gestational Age Infants 1983 and The Leiden Follow-Up Project on Prematurity 1996-1997. *Pediatrics*, *115*(2), 396–405.

<https://doi.org/10.1542/peds.2004-1497>

- Stoll, B., Ansen, E., Anaroff, V., Right, L., Carlo, W., Hrenkranz, R., ... Poole, K. (2002). Changes in Pathogens Causing Early-Onset Sepsis in Very-Low-Birth-Weight Infants. *The New England Journal of Medicine*, *58*(1), 17–18. <https://doi.org/10.1097/00006254-200301000-00010>
- Stoll, B., Hansen, N., Bell, E., Shankaran, S., Laptook, A., Walsh, M., ... Higgi, R. (2010). Neonatal outcomes of extremely preterm infants from the NICHD neonatal research network. *Pediatrics*, *126*(3), 443–456. <https://doi.org/10.1542/peds.2009-2959>. Neonatal
- Stoll, B. J. (1997). The Global Impact of Neonatal Infection. *Clinics in Perinatology*, *24*(1), 1–21. [https://doi.org/10.1016/s0095-5108\(18\)30181-7](https://doi.org/10.1016/s0095-5108(18)30181-7)
- Stoll, B. J. (2004). Neurodevelopmental and Growth Impairment Among Extremely Low-Birth-Weight Infants With Neonatal Infection. *Jama*, *292*(19), 2357. <https://doi.org/10.1001/jama.292.19.2357>
- Stoll, B. J., Hansen, N., Fanaroff, A. A., Wright, L. L., Carlo, W. A., Ehrenkranz, R. A., ... Poole, W. K. (2004). Late-Onset Sepsis in Very Low Birth Weight Neonates: The Experience of the NICHD Neonatal Research Network. *Pediatrics*, *110*(2), 285–291. <https://doi.org/10.1542/peds.110.2.285>
- Stoll, B. J., Hansen, N. I., Adams-Chapman, I., Fanaroff, A. A., Hintz, S. R., Vohr, B., ... Network, for the N. I. of C. H. and Human D. N. R. (2004). Neurodevelopmental and Growth Impairment Among Extremely Low-Birth-Weight Infants With Neonatal Infection. *Jama*, *292*(19), 2357. <https://doi.org/10.1001/jama.292.19.2357>
- Stoll, B. J., Hansen, N. I., Bell, E. F., Walsh, M. C., Carlo, W. A., Shankaran, S., ... Higgins, R. D. (2015). Trends in care practices, morbidity, and mortality of extremely preterm Neonates, 1993-2012. *JAMA - Journal of the American Medical Association*, *314*(10), 1039–1051. <https://doi.org/10.1001/jama.2015.10244>
- Sweeney, J. K., & Blackburn, S. (2013). Neonatal physiological and behavioral stress during neurological assessment. *Journal of Perinatal and Neonatal Nursing*, *27*(3), 242–252. <https://doi.org/10.1097/JPN.0b013e31829dc329>
- Synnes, A., Luu, T. M., Moddemann, D., Church, P., Lee, D., Vincer, M., ... Lee, S. K. (2017). Determinants of developmental outcomes in a very preterm Canadian cohort. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *102*(3), F235–F234. <https://doi.org/10.1136/archdischild-2016-311228>
- Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., ... de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, *60*(1), 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
- Thelen, E. (1985). Developmental origins of motor coordination: leg movements in human infants. *Developmental Psychobiology*, *18*(1), 1–22. <https://doi.org/10.1002/dev.420180102>

- Thelen, Esther. (1995). Motor development. A new synthesis. *The American Psychologist*, 50, 79–95. <https://doi.org/10.1037/0003-066X.50.2.79>
- Thompson, A. M., & Bizzarro, M. J. (2008). Necrotizing Enterocolitis in Newborns. *Drugs*, 68(9), 1227–1238. <https://doi.org/10.2165/00003495-200868090-00004>
- Touwen, B. (1990). VARIABILITY AND STEREOTYPY OF SPONTANEOUS MOTILITY AS A PREDICTOR OF NEUROLOGICAL DEVELOPMENT OF PRETERM INFANTS. *Developmental Medicine & Child Neurology*, 32, 501–508.
- Tsai, M. L., Lien, R., Chiang, M. C., Hsu, J. F., Fu, R. H., Chu, S. M., ... Yang, P. H. (2012). Prevalence and morbidity of late preterm infants: Current status in a medical center of Northern Taiwan. *Pediatrics and Neonatology*, 53(3), 171–177. <https://doi.org/10.1016/j.pedneo.2012.04.003>
- Ustad, T., Evensen, K. A. I., Bertocelli, N., Frassoldati, R., & Ferrari, F. (2017). Validity of the General Movement Optimality List in Infants Born Preterm. *Pediatric Physical Therapy*, 29(4), 315–320. <https://doi.org/10.1097/PEP.0000000000000445>
- Valentin, T., Uhl, K., & Einspieler, C. (2005). The effectiveness of training in Prechtl's method on the qualitative assessment of general movements. *Early Human Development*, 81(7), 623–627. <https://doi.org/10.1016/j.earlhumdev.2005.04.003>
- Valle, S. C., Støen, R., Sæther, R., Jensenius, A. R., & Adde, L. (2015). Test–retest reliability of computer-based video analysis of general movements in healthy term-born infants. *Early Human Development*, 91(10), 555–558. <https://doi.org/10.1016/j.earlhumdev.2015.07.001>
- van der Ree, M., Tanis, J. C., Van Braeckel, K. N. J. A., Bos, A. F., & Roze, E. (2011). Functional impairments at school age of preterm born children with late-onset sepsis. *Early Human Development*, 87(12), 821–826. <https://doi.org/10.1016/j.earlhumdev.2011.06.008>
- van Iersel, P. a M., Bakker, S. C. M., Jonker, A. J. H., & Hadders-Algra, M. (2010). Does perinatal asphyxia contribute to neurological dysfunction in preterm infants? *Early Human Development*, 86(7), 457–461. <https://doi.org/10.1016/j.earlhumdev.2010.06.003>
- Vergnano, S., Sharland, M., Kazembe, P., Mwansambo, C., & Heath, P. T. (2005). Neonatal sepsis: An international perspective. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, 90(3), 220–224. <https://doi.org/10.1136/adc.2002.022863>
- Volpe, J. (1998). Brain injury in the premature infant: Overview of clinical aspects, neuropathology, and pathogenesis. *Seminars in Pediatric Neurology*, 5(3), 135–151. Retrieved from <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1998317666>
- Volpe, J. (2009). The Encephalopathy of Prematurity – Brain Injury and Impaired

- Brain Development Inextricably Intertwined. *Semin Pediatr Neurol*, 16(4), 167–178. <https://doi.org/10.1016/j.spen.2009.09.005>.The
- Volpe, J. J. (1997). Brain injury in the premature infant - from pathogenesis to prevention. *Brain Dev*, 19(8), 519–534. [https://doi.org/10.1016/s0387-7604\(97\)00078-8](https://doi.org/10.1016/s0387-7604(97)00078-8)
- Volpe, Joseph. (2009). Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances Joseph. *Lancet Neurol*, 50(4), 309–312. [https://doi.org/10.1016/S1474-4422\(08\)70294-1](https://doi.org/10.1016/S1474-4422(08)70294-1).Brain
- Volpe, Joseph J. (2005). Encephalopathy of prematurity includes neuronal abnormalities. *Pediatrics*, 116(1), 221–225. <https://doi.org/10.1542/peds.2005-0191>
- Walsh, J. M., Doyle, L. W., Anderson, P. J., Lee, K. J., & Cheong, J. L. Y. (2014). Moderate and Late Preterm Birth: Effect on Brain Size and Maturation at Term-Equivalent Age. *Radiology*, 273(1), 232–240. <https://doi.org/10.1148/radiol.14132410>
- Watson, P. F., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. *Theriogenology*, 73(9), 1167–1179. <https://doi.org/10.1016/j.theriogenology.2010.01.003>
- Weir, J. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231–240.
- Weisglas-Kuperus, N., Baerts, W., Fetter, W. P. F., & Sauer, P. J. J. (1992). Neonatal cerebral ultrasound, neonatal neurology and perinatal conditions as predictors of neurodevelopmental outcome in very low birthweight infants. *Early Human Development*, 31(2), 131–148. [https://doi.org/10.1016/0378-3782\(92\)90040-N](https://doi.org/10.1016/0378-3782(92)90040-N)
- Whitley, E., & Ball, J. (2002). Statistics review 3: hypothesis testing and P values. *Critical Care (London, England)*, 6(3), 222–225. <https://doi.org/10.1186/cc1868>
- Williams, J. E., & Pugh, Y. (2018). The Late Preterm. *Critical Care Nursing Clinics of North America*, 30(4), 431–443. <https://doi.org/10.1016/j.cnc.2018.07.001>
- Wilson-Costello, D., Friedman, H., Minich, N., Fanaroff, A. A., & Hack, M. (2005). Improved Survival Rates With Increased Neurodevelopmental Disability for Extremely Low Birth Weight Infants in the 1990s. *Pediatrics*, 115(4), 997–1003. <https://doi.org/10.1542/peds.2004-0221>
- Woodward, L. J., Anderson, P. J., Austin, N. C., Howard, K., & Inder, T. E. (2014). Neonatal MRI to Predict Neurodevelopmental Outcomes in Preterm Infants, 685–694.
- Woythaler, M., McCormick, M. C., Mao, W., & Smith, V. C. (2015). Late Preterm Infants and Neurodevelopmental Outcomes at Kindergarten, 136(3). <https://doi.org/10.1542/peds.2014-4043>
- Wusthoff, C. J. (2013). How to use: the neonatal neurological examination. *Archives of Disease in Childhood. Education and Practice Edition*, 98(4), 148–153.

<https://doi.org/10.1136/archdischild-2013-303640>

- Wynn, J. L. (2016). Defining Neonatal Sepsis James. *Curr Opin Pediatr*, 28(2), 135–140. <https://doi.org/10.1097/MOP.0000000000000315>.Defining
- Wynn, J. L., Wong, H. R., Shanley, T. P., Bizzarro, M. J., Saiman, L., & Polin, R. A. (2011). Time for a neonatal–specific consensus definition for sepsis James. *The Pediatric Infectious Disease Journal*, 36(1), 24. <https://doi.org/10.1097/PCC.0000000000000157>.Time
- Yeh, K.-K., Liu, W.-Y., Wong, A. M.-K., Chung, C.-Y., Lien, R., & Chuang, Y.-F. (2016a). Intra-observer reliability of Prechtl's method for the qualitative assessment of general movements in Taiwanese infants. *Journal of Physical Therapy Science*, 28(5), 1588–1594. <https://doi.org/10.1589/jpts.28.1588>
- Yeh, K.-K., Liu, W.-Y., Wong, A. M.-K., Chung, C.-Y., Lien, R., & Chuang, Y.-F. (2016b). Intra-observer reliability of Prechtl's method for the qualitative assessment of general movements in Taiwanese infants. *Journal of Physical Therapy Science*, 28(5), 1588–1594. <https://doi.org/10.1589/jpts.28.1588>
- Yeo, K. T., Safi, N., Wang, Y. A., Marsney, R. Le, Schindler, T., Bolisetty, S., ... Lui, K. (2017). Prediction of outcomes of extremely low gestational age newborns in Australia and New Zealand, 1–8. <https://doi.org/10.1136/bmjpo-2017-000205>
- Younge, N., Goldstein, R., Bann, C., Hintz, S., Patel, R., Smith, B., ... Cotten, M. (2017). Survival and Neurodevelopmental Outcomes among Periviable Infants. *The New England Journal of Medicine*, 42(5), 301. <https://doi.org/10.1097/NMC.0000000000000360>
- Zahed-Cheikh, M., Brévaut-Malaty, V., Busuttill, M., Monnier, A.-S., Roussel, M., & Gire, C. (2011). Comparative analysis of perinatal and postnatal factors, and general movement in extremely preterm infants. *Brain and Development*, 33(8), 656–665. <https://doi.org/10.1016/j.braindev.2010.10.023>
- Zhong, G., Shevtsova, N. A., Rybak, I. A., & Harris-Warrick, R. M. (2012). Neuronal activity in the isolated mouse spinal cord during spontaneous deletions in fictive locomotion: Insights into locomotor central pattern generator organization. *Journal of Physiology*, 590(19), 4735–4759. <https://doi.org/10.1113/jphysiol.2012.240895>
- Zorzenon, R. F. M., Takaara, L. K., & Linhares, M. B. M. (2019). General spontaneous movements in preterm infants differentiated by post-conceptual ages. *Early Human Development*, 134(April), 1–6. <https://doi.org/10.1016/j.earlhumdev.2019.04.011>

Liste de figures

Figure 1. Variabilité des GMs	41
Figure 2. Variabilité réduite des GMs.....	41
Figure 3. Système d'analyse des GMs sur détection en 3D.....	47
Figure 4. Système d'analyse des GMs sur détection en 2D.....	48
Figure 5. Accéléromètre et système de coordonnées	49
Figure 6. Validité et fiabilité des mesures cliniques	54
Figure 7. Représentation de la concordance entre deux mesures répétées	59
Figure 8. Recrutement et évaluation des GMs	73
Figure 9. Diagnostic des GMs	79
Figure 10. Pourcentage de classifications des GMs dans les trois périodes	82
Figure 11. Score optimal des GMs à préterme et terme.....	84
Figure 12. Score optimal des GMs typiques et GMs atypiques à préterme et terme	85
Figure 13. Score optimal des classifications des GMs à préterme et terme	86
Figure 14. Score optimal des GMs à post terme	90
Figure 15. Score optimal des classifications des GMs à post terme.....	91
Figure 16. Fiabilité intracodeur pour le score optimal des GMs	109
Figure 17. Fiabilité intercodeur pour le score optimal des GMs	111

Liste de tableaux

Tableau 1. Mesures neuromotrices pour l'enfant né prématurément.....	30
Tableau 2. Premiers mouvements au cours du premier trimestre de gestation.....	39
Tableau 3. Tableau de décompte pour calculer des indices épidémiologiques.....	57
Tableau 4. Les états comportementaux du nouveau-né	75
Tableau 5. Diagnostic des GMs à préterme et terme	80
Tableau 6. Diagnostic des GMs à préterme et post terme	81
Tableau 7. Classifications des GMs à préterme et terme.....	83
Tableau 8. Classifications des GMs à préterme et post-terme.....	83
Tableau 9. Pourcentage de GMs remplissant les critères optimaux à préterme	87
Tableau 10. Pourcentage de GMs remplissant les critères optimaux à terme.....	89
Tableau 11. Pourcentage de GMs remplissant les critères optimaux à post terme.....	92
Tableau 12. Caractéristiques de l'échantillon	100
Tableau 13. Régression logistique pour le diagnostic des GMs.....	104
Tableau 14. Régression linéaire pour le score optimal des GMs	105
Tableau 15. Diagnostic des GMs et l'hémorragie intraventriculaire	106
Tableau 16. Diagnostic des GMs et la leucomalacie périventriculaire	107
Tableau 17. Régression linéaire pour le score des GMs et l'HIV ou la LPV	107
Tableau 18. Pourcentage d'accord intracodeur pour le diagnostic des GMs.....	108
Tableau 19. Pourcentage d'accord intercodeur pour le diagnostic des GMs.....	108
Tableau 20. Accord intracodeur pour les critères optimaux des GMs.....	110
Tableau 21. Accord intercodeur pour les critères optimaux des GMs.....	112

Liste d'encadrés


Encadré 1. Résumé du chapitre I.....	26
Encadré 2. Résumé du chapitre II.....	37
Encadré 3. Résumé du chapitre III.....	53
Encadré 4. Résumé du chapitre IV	67
Encadré 5. Résumé du chapitre V	70
Encadré 6. Résumé du chapitre VI	96
Encadré 7. Résumé du chapitre VII	117

Liste d'abréviations

OMS	Organisation mondiale de la santé
SA	Semaines d'aménorrhée
AG	Âge gestationnel
AC	Âge chronologique
Agc	Âge gestationnel corrigé
SPT	Semaines post-terme
g	Grames
IRM	Imagerie par Résonance Magnétique
DBP	Dysplasie bronchopulmonaire
EUNC	Entérocolite ulcéro-nécrosante
HIV	Hémorragie intraventriculaire
LPV	Leucomalacie périventriculaire
TIMP	Test of Infant Motor Performance
GMs	Mouvements généraux
TDS	Théorie de la dynamique des systèmes
NGST	Théorie de la sélection des groupes neuronaux
CPGs	Générateurs centraux de patterns
3D	Trois dimensions
2D	Deux dimensions
VPP	Valeur prédictive positive
VPN	Valeur prédictive négative
ICC	Le coefficient de corrélation intraclasse
IC	Intervalle de confiance
OR	Odds ratio
SD	Standard déviation (écart type)
IC BCa	Intervalle de confiance sur la méthode bootstrap pour corriger le biais
Méd	Médiane
min	Minimum
max	Maximum

Annexe 1

Feuille de codage pour l'évaluation détaillée des GMs à la période préterme et terme



The GM Trust

Detailed Assessment of General Movements (GMs) During Preterm and Term Age

Name _____		Date of Birth _____	
Recording Date _____		Postmenstrual Age _____ weeks	
Behavioural State (Coincidence) _____		<input type="checkbox"/> State 2 (Active Sleep) <input type="checkbox"/> State 4 (Active Wakefulness)	

<input type="checkbox"/> Normal	Sequence	<input type="checkbox"/> 2 variable
<input type="checkbox"/> Poor Repertoire		<input type="checkbox"/> 1 monotonous and/or broken
<input type="checkbox"/> Cramped-Synchronised		<input type="checkbox"/> 0 synchronised
<input type="checkbox"/> Hypokinetic		<input type="checkbox"/> 0 disorganised
<input type="checkbox"/> Chaotic		

Detailed Scoring

	Neck	Trunk
	<input type="checkbox"/> 2 involved in the sequence <input type="checkbox"/> 1 hardly or not involved	<input type="checkbox"/> 2 fluent and elegant rotations <input type="checkbox"/> 1 just a few rotations <input type="checkbox"/> 0 almost no rotations
	Upper Extremities	Lower Extremities
Amplitude	<input type="checkbox"/> 2 variable, full range <input type="checkbox"/> 1 predominantly small range <input type="checkbox"/> 1 predominantly large range <input type="checkbox"/> 1 neither small nor large but monotonous	<input type="checkbox"/> 2 variable, full range <input type="checkbox"/> 1 predominantly small range <input type="checkbox"/> 1 predominantly large range <input type="checkbox"/> 1 neither small nor large but monotonous
Speed	<input type="checkbox"/> 2 variable <input type="checkbox"/> 1 monotonously slow <input type="checkbox"/> 1 monotonously fast <input type="checkbox"/> 1 neither small nor fast but monotonous	<input type="checkbox"/> 2 variable <input type="checkbox"/> 1 monotonously slow <input type="checkbox"/> 1 monotonously fast <input type="checkbox"/> 1 neither small nor fast but monotonous
Spatial range	<input type="checkbox"/> 2 full space variably used <input type="checkbox"/> 1 limited space <input type="checkbox"/> 0 in one plane only	<input type="checkbox"/> 2 full space variably used <input type="checkbox"/> 1 limited space <input type="checkbox"/> 0 lifted-released
Proximal rotatory components	<input type="checkbox"/> 2 present, fluent and elegant <input type="checkbox"/> 1 just a few rotations <input type="checkbox"/> 0 almost no rotations	<input type="checkbox"/> 2 present, fluent and elegant <input type="checkbox"/> 1 just a few rotations <input type="checkbox"/> 0 almost no rotations
Distal rotatory components	<input type="checkbox"/> 2 present, fluent and elegant <input type="checkbox"/> 1 just a few rotations <input type="checkbox"/> 0 almost no rotations	<input type="checkbox"/> 2 present, fluent and elegant <input type="checkbox"/> 1 just a few rotations <input type="checkbox"/> 0 almost no rotations
Onset	<input type="checkbox"/> 2 smooth and fluctuating <input type="checkbox"/> 1 minimal fluctuations <input type="checkbox"/> 0 predominantly abrupt	<input type="checkbox"/> 2 smooth and fluctuating <input type="checkbox"/> 1 minimal fluctuations <input type="checkbox"/> 0 predominantly abrupt
Offset	<input type="checkbox"/> 2 smooth and fluctuating <input type="checkbox"/> 1 minimal fluctuations <input type="checkbox"/> 0 predominantly sudden release	<input type="checkbox"/> 2 smooth and fluctuating <input type="checkbox"/> 1 minimal fluctuations <input type="checkbox"/> 0 predominantly sudden release
Tremulous movements	<input type="checkbox"/> 2 absent <input type="checkbox"/> 1 unilaterally present <input type="checkbox"/> 0 bilaterally present	<input type="checkbox"/> 2 absent <input type="checkbox"/> 1 unilaterally present <input type="checkbox"/> 0 bilaterally present
Cramped components	<input type="checkbox"/> 2 absent <input type="checkbox"/> 1 occasionally present <input type="checkbox"/> 0 predominantly present	<input type="checkbox"/> 2 absent <input type="checkbox"/> 1 occasionally present <input type="checkbox"/> 0 predominantly present


Optimality subscores	Upper extremities (max 18) <input style="width: 50px;" type="text"/>	Lower extremities (max 18) <input style="width: 50px;" type="text"/>	
	Neck and Trunk (max 4) <input style="width: 50px;" type="text"/>	Sequence (max 2) <input style="width: 50px;" type="text"/>	
GM Optimality Score (max. 42) <input style="width: 50px;" type="text"/>			

Annexe 2

Feuille de codage pour l'évaluation détaillée des GMs à la période post terme

Assessment of Motor Repertoire - 3 to 5 Months

Christa Einspieler and Arie Bos, the GM Trust 2000



Ref: Einspieler C, Prechtl HFR, Bos AF, Ferrari F, Cioni G. Prechtl's Method of the Qualitative Assessment of General Movements in Preterm, Term and Young Infants. Clin Dev Med 167. London: MacKeith Press 2004, page 26.

Name:

born: PMA: BW:

Recording Date: Age:

Observed movement patterns:

Number (excl.fidgety mov.) ▣ normal ▣ abnormal

N A fidgety movements		
N A swipes	N A hand-mouth contact	N A arching
N A wiggling-oscillating	N A hand-hand contact	N A rolling to side
N A kicking	N A fiddling / cloths, blanket	N A visual scanning
N A excitement bursts	N reaching	N hand regard
N A smiles	N A foot-foot contact	N A head antelexion
N A mouth movements	A asym segm movements	A arm movements in circles
A tongue movements	N A legs lift	A absent leg movements
N A head rotation	N A hand-knee contact	A

Observed postural patterns: ▣ normal ▣ abnormal

N A head in midline (20 °)	N variable finger postures	A hyperextension of the neck
N A symmetrical	A predominant fisting	A hyperextension of trunk
N A spontaneous ATNR absent or could be overcome	A finger spreading	A extended arms / on / above surface
A body and limbs 'flat' on surface	A few finger postures	A extended legs
	A synchronised opening and closing	

Movement character (global score):

N smooth and fluent	A stiff	A predominantly slow speed
A jerky	A cramped-synchronized	A predominantly fast speed
A monotonous		A predominantly large ampl.
A tremulous		A predominantly small ampl.

Motor Optimality List:

1.	Fidgety Movements	normal	<input type="checkbox"/>	12
		abnormal	<input type="checkbox"/>	4
	± + ++ * **	absent / sporadic	<input type="checkbox"/>	1
2.	Repertoire of co-existent other movements	age-adequate	<input type="checkbox"/>	4
		reduced	<input type="checkbox"/>	2
		absent	<input type="checkbox"/>	1
3.	Quality of other movements	N > A	<input type="checkbox"/>	4
		N = A	<input type="checkbox"/>	2
		N < A	<input type="checkbox"/>	1
4.	Posture	N > A	<input type="checkbox"/>	4
		N = A	<input type="checkbox"/>	2
		N < A	<input type="checkbox"/>	1
5.	Movement character	smooth and fluent	<input type="checkbox"/>	4
		abnormal but not CS	<input type="checkbox"/>	2
		cramped-synchronized	<input type="checkbox"/>	1

Motor Optimality Score:

from 28 to 5

Annexe 3

General Movements Assessment Validity on Preterm Infants: A-systematic Review

Angélica Valencia-Diaz^a, Michèle Molina^b, François Jouen^a, Joëlle Provasi^a

Affiliations: ^aCHArt, Ecole Pratique des Hautes Etudes, Paris Sciences Lettres, France,
^bLPCN, Université de Caen, France

Address correspondence to: Angélica Valencia, Ecole Pratique des Hautes Etudes, Les Patios Saint-Jacques, 4-14 Rue Ferrus, 75014 Paris, angelica.valencia@etu.ephe.fr

ABSTRACT

AIM: To find and evaluate evidence on discriminative and predictive validity of the General Movement Assessment (GMA) to assess neurologic functioning on preterm infants.

METHOD: Search of evidence was conducted with electronic databases up to February 2018. Screening and inclusion of studies were based on the following eligibility criteria: infants with gestational age (GA) of <37wk; initial GMA at preterm period; evaluating of GMA relationship with neonatal adversities; and outcome neurodevelopmental. Standardized forms were used to assess studies and extract information about participants, assessment, outcome, and statistical indexes.

RESULTS: Twenty-seven studies met eligibility criteria, involving 1340 participants with GA from 23wk to 36wk differing in clinical background. Discriminative validity evidence showed that the GMA was significantly affected by the following neonatal adversities: GA of ≤ 32 wk, complications, and brain lesions. Predictive validity evidence indicated that the GMA sensitivity ranged from 80% to 100% for adverse outcomes, and its specificity from 8% to 83% for normal outcome.

INTERPRETATION: The GMA had good discriminative and predictive validity on higher-risk preterm infants to detect neurological dysfunction. It identifies infants who might need more comprehensive developmental evaluation. The GMA is a suitable tool to screening purpose. That need confirming by validation studies.

INTRODUCTION

Nearly 7% of preterm infants will face neurological impairments that may lead to a lifelong need for additional care¹. Burden of disability might be reduced by early intervention. In fact, one literature review² concluded that preterm infants benefit more from developmental interventions beginning in neonatal intensive care units (NICU). Early intervention depends on neurologic evaluation to identify infants with neurological dysfunction, and those at risk for neurodevelopmental disability^{3,4}. Neurologic evaluation of preterm infants should be based on assessments appropriate for them.

The choice of an assessment instrument requires considering its construct, purposes, and measurement properties. A construct is a well-defined subject of measurement, in this case neurological function. The purpose of neurological assessment is threefold: discriminative to distinguish infants with a neurological dysfunction from those with a typical function; predictive to classify infants according to their developmental prognosis; and evaluative to measure changes in neurological status. Measurement properties imply validity, namely the degree to which an instrument measures the construct it purports to measure. Criteria of

discriminative validity is construct validity, that is the relation of the score of a measurement instrument with neonatal adversities impacting neurologic function. Criteria of predictive validity is criterion validity, that is the relation of the score of a measurement instrument with the gold standard in the future or outcome neurodevelopmental⁵.

The General Movements Assessment (GMA) is an instrument aimed at assessing the functional integrity of a young nervous system. Designed to be used from preterm age, it is based on observation of an age-related motor repertoire named General Movements (GM). Complexity, variability and fluidity of the GM reflect the normal neurological function. The latter is considered as abnormal when the GM characteristics are reduced⁶. A review⁷ of assessments for preterm infants during the first year of life shows that the GMA presented the best clinical utility, evaluative validity, and reliability ($k > 0.85$) compared to the other assessments reviewed. Three reviews⁸⁻¹⁰ on GMA measurement properties found that it had a highly predictive validity (sensitivity ≥ 0.92 ; specificity ≥ 0.82) on heterogeneous population.

Although validity of an assessment instrument might vary depending on purpose, population tested, and setting, no systematic review has specifically focused on the validity of the GMA performed during the preterm period. To guide early clinical intervention and future researches, this systematic review is aimed at finding out and evaluating published evidence on discriminative and predictive validity of the GMA performed on preterm infants during the preterm period.

METHOD

Search Strategy

Three comprehensively literature searches were conducted with the following electronic databases: MEDLINE (1979 to February 2018), Embase (1977 to February 2018), Science Direct (1979 to February 2018), Cinahl (1982 to February 2018), Web of Science (1967 to February 2018). The keyterms used for the first search were 'preterm infant' or 'low birthweight infant', and 'neurological function' or 'neurological dysfunction' or 'brain function', and 'general movements' or 'general movements assessment', and 'discriminative validity'. The keyterms for the second search were 'preterm infant' or 'low birthweight infant', and 'child development' or 'neurodevelopmental outcome' or

‘neurodevelopmental disabilities’, and ‘general movements’ or ‘general movements assessment’, and ‘predictive validity’. The terms ‘discriminative’, ‘predictive’ and ‘validity’ were suppressed, because they lead to too few results (9 items). A third search was conducted for each name author: Bos AF., Cioni G., Einspieler C., Hadders-Algra M., and Prechtl HF. The fourth search was conducted on the General Movements Trust website. The electronic search was supplemented by a manual search of the reference lists of the selected studies.

Study Selection

All publication status and study designs were considered. Studies were included if they fulfilled the following criteria: (1) participants were preterm infants with gestational age (GA) of <37wk; (2) the initial GMA was performed at preterm period and its analysis not pooled with analyses conducted at other periods, because the GM evolve from the term age to 5mo⁶; (3) the relation between the GMA and neonatal adversities impacting neurological function was evaluated or groups varying risk levels were compared; (4) the relation between the GMA and outcome according to neurodevelopmental assessment was evaluated. Eligibility criteria were applied to titles and abstracts of the search results, and full text of potentially relevant studies was retrieved to confirm selection. Studies were excluded if they were: (1) not published in English, French or Spanish language; (2) not a research report; (3) not based on the Prechtl’s method for the GMA. Figure 1 details study selection.

Extraction and Assessment Process

Data extracted from included studies regarded sampling, characteristics of participants, age at the GMA, outcome measure, follow-up age, and statistical indexes. The *P* value determining statistical significance of evaluated relations was extracted. Indexes of correlation such as Spearman’s coefficient (*r*_s) were also considered. Measures evaluating accuracy of diagnostic tests were extracted. The accuracy measures were: sensitivity corresponding to the proportion of true-positives correctly identified; and specificity corresponding to the proportion of true-negatives correctly identified. Extraction took

account of population estimates such as confidence intervals (CI) that express the range of values that is plausible for the population.

The quality evidence was assessed by a modified version⁸ of the Critical Review Form-Quantitative Studies¹¹ containing 12 items. Items were scored with 1 or 0 if criteria were met or not, and the sum of them produced a global score rating the methodological quality of the included studies as very high (>80%), high (>60%-80%), and low (<60%). Extraction and assessment were done by AV and confirmed by JP and FJ using standardized forms based on criteria for studies on test accuracy¹²⁻¹⁴.

RESULTS

Study Characteristics and Quality Assessment

Of the 2628 references that were found, 820 were screened to determine their eligibility. Twenty-seven observational studies¹⁵⁻⁴¹ fulfilling the eligibility criteria were included in this review. Included studies involved 1340 participants whose mean GA varied from 23wk to 36wk, and mean birth weight (BW) from 490g to 2750g. All studies included both participants with brain abnormality and participants without it according to ultrasonography (US). All studies used the GMA in a research setting, the most of which reported initial assessment in NICU. Few studies have specified weeks of post-menstrual age (PMA) and time points of the GMA. Outcome measures being reported are magnetic resonance imaging (MRI), and diverse neurodevelopmental assessments at follow-up ages between 12mo corrected age (CA) and 10yr. Studies conveyed results using statistical indexes, except one study²² presenting a cross-classification table from which we calculated diagnostic accuracy measures. Table 1 presents characteristics of included studies.

About methodological quality, studies obtained a mean score of 9.3 according to the Critical Review Form – Quantitative Studies. The most problematic item was avoiding risk of bias in the following aspects: formulating hypothesis^{21-25,27,29-34,37,41}, sampling procedure^{16-18,20-23,26,27,29,30,33,34,38,41}, assessment procedure^{19,20,22,23,26,32,40}, reliability testing^{15,19-21,23,25,28,38}, and data reporting^{15-23,25-27,29-34,38}. Table 2 details assessment of quality of included studies.

Evidence from studies was narratively synthesized. Studies on the GMA discriminative validity were grouped according to relation of the GMA with neonatal adversities or

differences between groups. Studies on the predictive validity of the GMA were pooled according to follow-up age of neurodevelopmental assessments.

Discriminative Validity of the General Movements Assessment

Concerning relation of the GMA with GA, it appears that ≤ 32 wk was a key GA. This relation was evaluated in nine studies^{20,21,26,31,32,35,38,40,41}, one of which found increasing abnormal GM in those youngest of a sample of babies of < 32 wk GA³⁸. This outcome was similar in two samples differing in risk, as abnormal GM were related to GA of ≤ 32 wk in infants without complications²⁶, and infants with perinatal asphyxia⁴¹. In contrast, decreasing odds of abnormal GM was related to increasing of GA in those youngest (with 32wk of GA) of a sample of infants with ≥ 32 wk GA⁴⁰. A better GM trajectory was correlated ($r_s=0.418$) with highest GA in infants of ≤ 32 wk GA²¹, and highest GM score was associated with highest GA in extremely low birth weight (ELBW) infants²⁰. Also, there was conflicting evidence about relation of the GMA with GA, since one study showed that GM quality related to GA³¹ in a sample of infants of < 36 wk GA, while another study found that the GMA was not related to GA in infants of ≤ 33 wk GA³².

Evidence about relation of the GMA with BW was conflicting according to reports from seven studies^{16,20,21,26,32,35,38}. In this way, one study found that abnormal GM were related to lower BW in infants of < 32 wk GA³⁸, and one study showed that normal GM correlated ($r_s=0.53$) with higher BW in infants of ≤ 32 wk GA²¹. On the contrary, one study reported that GM quality was not related to BW on infants with ≤ 33 wk GA³², and one study found that GM quality was not related to BWz score in infants of ≤ 30 wk GA³⁵. Furthermore, two studies found that abnormal GM was related to small for gestational age (SGA) in SGA infants¹⁶, and infants without medical complications²⁶, whilst one study found that GM quality was not related to SGA, nor appropriate for gestational age (AGA) in infants of < 32 wk GA³⁸. Surprisingly, one study found that higher GM scores correlated with SGA in ELBW infants²⁰.

The relation between the GMA and clinical events depends on the nature of the latter. This relation was assessed in six studies^{20,32,35,38,39,41}. Three of these studies showed that lower GM scores and abnormal GM related to sepsis^{20,35}, necrotizing enterocolitis³⁵, bronchopulmonary dysplasia³⁵, respiratory problems⁴¹, and hyperbilirubinemia²⁰.

Nevertheless, two studies found that GM quality was not related to placental pathology³⁸, neither perinatal asphyxia⁴¹. Although the GMA and AV (Artificial Ventilation) were related in two studies^{20,38}, GM quality and AV days were not related in two studies^{32,39}.

Evidence seems to show that relation of the GMA with brain image depends on brain image findings. This relation was evaluated in eleven studies^{20–23,25,28,29,31,33,36,38}. Two studies found that abnormal GM were related to cortical grey matter abnormality, larger interhemispheric distance, and abnormality of cerebellum, but not with other brain abnormalities detected on MRI in both infants of ≤ 30 wk GA³⁶ and < 31 wk GA²⁸. Also, two studies demonstrated that abnormal GM were related to brain injury severity according to US on both infants with motor abnormalities²⁹, and infants of < 36 wk GA³¹. In fact, three studies presenting accuracy measures showed that the sensitivity of the GMA ranged from 71% to 100%, and its specificity from 42% to 92.8%^{22,23,25}. Both lowest sensitivity (71%) and lowest specificity (42%) were found in the same sample of infants for which periventricular hemorrhage grades 1–3 was the usual cerebral abnormality²⁵. The highest sensitivity (100%) and highest specificity (92.8%) were obtained in one study on infants presenting principally lesions of the periventricular white matter²³. However, there was conflicting evidence about relation between the GMA and brain image, because two studies showed that the GMA was related to US in ELBW²⁰, and LBW infants³³, whereas three studies found no significant relation between GM quality and US in infants of < 35 wk GA^{21,25,38}.

Evidence on the GMA differences between high and low-risk infants was conflicting. This was addressed in four studies^{18,26,31,41} by comparing GM quality between groups that varied in neurological risk. Three studies found that abnormal GM were more observed in high-risk groups, which included infants with brain injury¹⁸, high perinatal risk³¹, and SGA infants²⁶. Conversely, abnormal GM were less observed in low-risk groups, including infants with non-severe brain injury¹⁸, low perinatal risk³¹, and AGA infants²⁶. However, one study found that GM quality did not differ between infants with perinatal asphyxia and those without it⁴¹.

Predictive Validity of the General Movements Assessment

Relation of the GMA with neurodevelopmental assessments at 12mo CA to 23mo CA showed that accuracy predictive of the GMA was high for predicting abnormal outcomes, but unstable for normal one. This was examined in eight studies^{15,19,20,22,25,27,32,37}. These studies found that sensitivity values ranged between 40% and 100%. The lowest value (40%) reflected the GMA sensitivity for borderline outcome according to the Touwen Infant Neurological Evaluation (TINE) on infants without medical complications²⁷. The highest value (100%) showed the GMA sensitivity for both definite focal abnormalities by neurological examination, and abnormal profile by the Denver Developmental Screening Test (DDST) on infants of <35wk GA²⁵, and infants with brain injury²². Concerning specificity, values fluctuated between 8% and 83%. The lowest value (8%) corresponded to the specificity of the GMA, performed at 34-36wk PMA, for normal development according to the Alberta Infant Motor Scales (AIMS) on infants of <30wk GA³⁷. The highest value (83%) reflected the specificity of the GMA, performed between 30-36wk PMA, for normal development according to neurological examination on infants of <32wk GA¹⁵. Evidence on the GMA predictive validity based on univariate and regression analysis was conflicting. While two studies found that the GMA was related to neurological examination¹⁵, the AIMS, and the Paediatric Evaluation of Disability Inventory (PEDI) ($r_s=0.41$)³², three studies reported no relation between the GMA and neurological examination, the AIMS, the DDST, and the Neurological, Sensory, Motor, Developmental Assessment^{20,25,37}.

Association between the GMA and neurodevelopmental assessments at 2yr to 3yr demonstrated that the predictive accuracy of the GMA was variable, being higher for abnormal than normal outcomes. This association was evaluated in five studies^{17,21,23,24,41}. About sensitivity, values were between 80% and 100%. The lowest value (80%) reflected the GMA sensitivity for minor abnormality according to paediatric evaluation and the TINE on infants of ≤ 32 wk GA²¹. The highest value (100%) corresponded to the GMA sensitivity for abnormal outcome, notably cerebral palsy (CP), according to the Amiel-Tison examination, the TINE, and the Griffiths scales on infants with brain injury^{23,24}. Concerning specificity, values were lower ranging between 38% and 59%. The lowest value (38%) indicated the GMA specificity for normal outcome based on developmental scales²⁴. The highest value (59%) corresponded to the GMA specificity for normal outcome according to

neurological examination and developmental scales²³. On the other hand, two studies performing univariate analysis found that the GMA was not related to neurological outcomes, neither CP according to the TINE on infants of ≤ 32 wk GA²¹, and infants with perinatal asphyxia⁴¹.

The relation between the GMA and neurodevelopmental assessments at 4yr to 10yr was examined in two studies^{30,34}. Through univariate analysis, these studies showed that the GMA was not related to neurodevelopmental outcomes according to the Hempel neurological examination, the TINE, and parental questionnaires on infants with abnormal GM.

DISCUSSION

This systematic review synthesizes and appraises evidence from 27 studies on the discriminative and predictive validity of the GMA on preterm infants during the preterm period. To our knowledge, this is the first review specifically focusing on measurement properties of the GMA during the preterm period. The following section summarizes the findings taking into account the quality of evidence for each main outcome. Thereafter, some methodological issues are addressed and limitations at review-level discussed.

Discriminative Validity of the General Movements Assessment

Most of the studies supported the discriminative validity of the GMA on preterm infants at high risk. That was proved by evidence-based on the relation of the GMA with neonatal adversities impacting neurologic functioning.

A good GMA discriminative validity was confirmed in extreme preterm infants (< 28 wk of GA) and very preterm infants (28wk GA to 32wk GA). Extremely and very preterm births increase risk of neurological impairment as GA decreases⁴². Impact of this risk factor on the GM was supported by high quality^{20,21,26,40,41} and very high-quality^{35,38} evidence showing that GM quality was more affected by the lower GA in the preterm infants of ≤ 32 weeks GA.

The GMA had a good discriminative validity in the sickest preterm infants⁴³⁻⁴⁵. This outcome was provided by very high-quality^{35,38,39} and high-quality^{20,41} evidence presenting that the GMA distinguish motor changes produced by infections, complications, and respiratory insufficiency, which involve neurological adversities in preterm infants⁴³⁻⁴⁵.

The discriminative validity of the GMA was good in preterm infants with more severe brain abnormalities. This outcome arises from high-quality evidence^{22,23,25,29,31} proving that the GMA detected motor abnormality relating to parenchymal lesions that affect neurological function⁴⁶. In fact, the GMA presented a high (>0.6) sensitivity, although an unstable specificity ranging between low (<0.6) and very high (>0.8). High sensitivity of a test involves a value to screening, but low specificity increases false-positive identification that requires confirming by complementary assessments¹². The GM can have a value for early identifying of preterm infants that could need more intensive diagnosis and assessment.

Some points could require further investigation to obtain a better perspective on the GMA discriminative validity. As test accuracy varies with the spectrum and prevalence of disease, it may be interesting to conduct more studies focusing on subgroups like moderate to late preterm infants, because they are increasing trend, and association with neurological dysfunction compared with term infants⁴⁷. High quality studies would be necessary to examine GMA differences between samples of premature infants which differ in neurological risk level. It would be interesting BW sampling-based studies to clarify relation between the GMA and LBW, because LBW includes heterogeneous groups of infants at risk: some are only preterm infants, but some are both preterm and SGA infants.

Predictive Validity of the General Movements Assessment

Most of the studies supported specific GMA validity for predicting the worst developmental outcomes in high-risk infants. However, there was also conflicting evidence on the GMA predictive validity. These outcomes were based on evaluation of relation between the GMA and diverse neurodevelopmental assessments.

The predictive validity of the GMA was good for severe abnormalities, but lower for minor disabilities and unstable for normal development. These outcomes were backed by high-quality^{15,17,19,21–23,25,27,37} and very high-quality²⁴ evidence, showing a very variable GMA predictive accuracy. In fact, the GMA had a very high sensitivity (>0.8) for predicting CP and lower developmental scores in infants at high-risk (as brain injured infants). In contrast, the sensitivity of the GMA was low (<0.6) to predict minor abnormalities (as borderline outcome) in low-risk infants. Additionally, the specificity of

the GMA was widely fluctuating between low (<0.6) and very high (>0.8) for normal development. Given its high sensitivity, but its low specificity, the GMA should be used as a complementary developmental assessment for screening and helping to identify infants who might need more comprehensive assessment and those that do not need it⁴⁸. In general, these outcomes indicate the variability of the predictive accuracy of the GMA, to the extent that it is a sensitive instrument for predicting abnormal developmental outcomes, but it is not specific to predicting the normal outcome. In fact, the variability of the predictive validity of the GMA was confirmed in two reviews^{9,49} on the GMA properties. Accordingly, the GMA had a good predictive validity for neurodevelopment in preterm infants under intensive care, but it was limited to minor disabilities in heterogeneous population.

The predictive validity of the GMA was not supported by high-quality^{20,21,25,30,37,41} and very high-quality evidence³⁴ that showed no significant relation between the GMA and neurological outcome. This outcome is confirmed in current literature showing no relation between the GMA and classic neurological assessments, discussing that the GMA evaluates a unique neurologic construct³⁹.

To improve interpretation of the GMA predictive validity, certain aspects can be considered. Long-term follow-up studies are necessary to evaluate the complementary role of GMA with other assessments for mild disabilities. Although GMA complementary role has been demonstrated to improve neuromotor prognosis³, this has been less examined about low-cognitive deficits, which are now the dominant sequelae in preterm infants. Future studies should clarify both PMA and time points of the GMA, because its accuracy was much less when it was performed at 34wk-36wk PMA compared to 30wk-36wk PMA in very preterm infants^{15,37}. Additionally, it has been suggested better prognosis with longitudinal GMA on heterogeneous populations.

Methodological Considerations

The quality of the evidence on the discriminative and predictive validity of the GMA on preterm infants during the preterm period was between high (70%) and very high (91%), since all studies included in this review met most of the criteria of the Critical Review Form-Quantitative Studies. As most of research on accuracy of tests^{13,14}, studies here reviewed presented specific problems that might potentially bias the results. First, few studies stated hypotheses about expected relations between the GMA and factors associated

with neurological function. This could make difficult to establish the GMA relevance for neurological assessment.

The most common problem was related to the sampling procedure. Some studies did not describe sampling procedures (i.e. inclusion by patient admission or selective convenience). Lacking sampling description makes it difficult to evaluate potential bias on both selection process and estimations on the studied population.

An important issue was associated with the assessment procedure. Most of studies not detailed content, classification criteria, and properties of instruments used. This makes it difficult to evaluate their similarity, dissimilarity and suitability to assess neurological functioning on population studied. Non-independent assessing represented a problem since awareness of GMA results, neurologic status, or participant clinical background could influence evaluators leading to overestimation of assessment diagnostic accuracy. Another risk of bias consisted of the lack of reliability testing description, given that based on observer judgement the GMA should be evaluated by measures of observer variability.

There were problems regarding data reporting. Most of the included studies not reported population estimates (i.e. CI) thus limiting the evaluation of the GMA clinical relevance. Estimate of the population provides useful information about both the magnitude of the evaluated factors and the stability of diagnostic accuracy measures. This was particularly true in small samples, because the level of uncertainty depends on the number of cases observed. Another issue was that few studies reported the manner how lost data were handled, limiting assessing the GMA clinical relevance due to the lack of clarity about differences in samples with a significant loss of data.

Limitations of the Systematic Review

A limitation was related to the selection process. We decided not to restrict selection to studies on the GMA validation since initial literature search including this topic gave few items. This could limit interpretation of cumulative evidence on the discriminative and predictive validity of the GMA, because most of the included studies were not designed to validate the GMA. Assessment process represented another limitation. The lack of appropriate methods for evaluating both health research⁴⁹ and studies on prognostic tests⁵⁰ is recognized. Therefore, we evaluated included studies on discriminative and predictive

validity using the same appraisal tool. We hence used the Critical Review Form-Quantitative Studies, because it is suitable for appraising reports on health research, and because there were some studies that presented both discriminative and predictive evidence on the validity of the GMA. This may not address important specific issues for each purpose, discriminative and predictive, of the GMA.

CONCLUSIONS

The present review indicates that the GMA had good discriminative validity during the preterm period to detect motor abnormalities related to neurological dysfunction. The GMA could be used for screening purpose on higher-risk preterm infants to identifying those who may need more intensive diagnosis assessments.

The GMA predictive validity was variable. The GMA accuracy was better for the prognosis of the worse developments than mild abnormal or normal development. Also, the GMA was highly accurate for predicting developmental outcome at 2yr to 3yr, when the CP diagnosis can be established.

Therefore, the GMA could be helpful as developmental screening assessment for identifying higher-risk preterm infants who may need more comprehensive evaluation for confirming the future course of disabilities or close follow-up.

Contradictory evidence on the discriminative and predictive validity of the GMA may reflect the lack of specifically designed studies that evaluate GMA measurement properties. More research on GMA validation during the preterm period, preferably, in a clinical setting, is necessary.

REFERENCES

1. Blencowe H, Lee AC, Cousens S, Bahalim A, Narwal R, Zhong N, et al. Preterm birth–associated neurodevelopmental impairment estimates at regional and global levels for 2010. *Pediatr Res* [Internet]. 2013;74(december):17–34. Available from: <http://www.nature.com/doi/10.1038/pr.2013.204>
2. Blauw-Hospers CH, de Graaf-Peters VB, Dirks T, Bos a. F, Hadders-Algra M. Does early intervention in infants at high risk for a developmental motor disorder improve motor and cognitive development? *Neurosci Biobehav Rev*. 2007;31(8):1201–12.
3. Hadders-Algra M. Early Diagnosis and Early Intervention in Cerebral Palsy. *Front Neurol* [Internet]. 2014;5(September):1–13. Available from: <http://journal.frontiersin.org/article/10.3389/fneur.2014.00185/abstract>
4. Majnemer A, Mazer B. Neurologic evaluation of the newborn infant: Definition and psychometric properties. *Dev Med Child Neurol* [Internet]. 1998;40:708–15. Available from: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed4&NEWS=N&AN=1998387524>
5. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38(1):27–36.
6. Einspieler C, Prechtl HFR. Prechtl’s assessment of general movements: A diagnostic tool for the functional assessment of the young nervous system. *Ment Retard Dev Disabil Res Rev*. 2005;11(1):61–7.
7. Spittle A, Doyle L, Boyd R. A systematic review of the clinimetric properties of neuromotor assessments for preterm infants during the first year of life. *Dev Med Child Neurol* [Internet]. 2008;50(4):254–66. Available from: <http://doi.wiley.com/10.1111/j.1469-8749.2008.02025.x>
8. Burger M, Louw Q a. The predictive validity of general movements – A systematic review. *Eur J Paediatr Neurol* [Internet]. 2009;13(5):408–20. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1090379808001839>
9. Darsaklis V, Snider LM, Majnemer A, Mazer B. Predictive validity of Prechtl’s Method on the Qualitative Assessment of General Movements: a systematic review of the evidence. *Dev Med Child Neurol* [Internet]. 2011;53(10):896–906. Available from: <http://doi.wiley.com/10.1111/j.1469-8749.2011.04017.x>
10. Einspieler C, Bos AF, Libertus ME, Marschik PB. The General Movement Assessment Helps Us to Identify Preterm Infants at Risk for Cognitive Dysfunction. *Front Psychol* [Internet]. 2016;7(March):406. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27047429>
11. Law, M. Stewart, D. Pollock, N. Letts, L. Bosch, J & Westmoreland M. Guidelines for Critical Review form- Quantitative Studies. Quant Rev Form-Guidelines [Internet]. 1998;1–11. Available from: <http://srs-mcmaster.ca/research/evidence-based-practice-research-group/>
12. Maxim LD, Niebo R, Utell MJ. Screening tests: a review with examples. *Inhal Toxicol* [Internet]. 2014;26(13):811–28. Available from: <http://www.tandfonline.com/doi/full/10.3109/08958378.2014.955932>
13. Carrington M, Lachs M, Feinstein A. Use of Methodological Standards in

- Diagnostic Test Research. *JAMA*. 1995;274:645–51.
14. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, Van der Meulen JHP, et al. Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. *Jama* [Internet]. 1999;282(11):1061–6. Available from: <http://dx.doi.org/10.1001/jama.1995.03520290060030>
<http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1995.03520290060030>
 15. Albers S JG. Prognostic Significance of Spontaneous Motility in Very Immature Preterm Infants under Intensive Care Treatment. *Biol Neonate*. 1994;
 16. Bos AF, van Loon AJ, Hadders-Algra M, Martijn A, Okken A, Prechtl HFR. Spontaneous motility in preterm, small-for-gestational age infants II. Qualitative aspects. *Early Hum Dev* [Internet]. 1997;50(1):131–47. Available from: <http://www.sciencedirect.com/science/article/pii/S0378378297000984>
 17. Cioni G, Ferrari F, Einspieler C, Paolicelli P, Barbani T, Prechtl H. Comparison between observation of spontaneous movements and neurologic examination in preterm infants. *J Pediatr* [Internet]. 1997;130(5):704–11. Available from: <http://www.sciencedirect.com/science/article/pii/S0022347697800108>
 18. Cioni G, Bos A, Einspieler C, Ferrari F, Martijn A, Paolicelli PB, et al. Early neurological signs in preterm infants with unilateral intraparenchymal echodensity. *Neuropediatrics* [Internet]. 2000;31(5):240–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11204281>
 19. Constantinou JC, Adamson-Macedo EN, Mirmiran M, Fleisher BE. Movement, imaging and neurobehavioral assessment as predictors of cerebral palsy in preterm infants. *J Perinatol* [Internet]. 2007;27(4):225–9. Available from: <http://www.nature.com/doi/10.1038/sj.jp.7211664>
 20. de Vries NKS, Erwich JJHM, Bos AF. General movements in the first fourteen days of life in extremely low birth weight (ELBW) infants. *Early Hum Dev*. 2008;84(11):763–8.
 21. de Vries NKS, Bos AF. The quality of general movements in the first ten days of life in preterm infants. *Early Hum Dev*. 2010;86(4):225–9.
 22. Droit S, Boldrini A, Cioni G. Rhythmical leg movements in low-risk and brain-damaged preterm infants. *Early Hum Dev*. 1996;44(3):201–13.
 23. Ferrari F, Cioni G, Prechtl HF. Qualitative changes of general movements in preterm infants with brain lesions. *Early Hum Dev* [Internet]. 1990;23(3):193–231. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2253580>
 24. Ferrari F, Cioni G, Einspieler C, Roversi MF, Bos AF, Paolicelli PB, et al. Cramped synchronized general movements in preterm infants as an early marker for cerebral palsy. *Arch Pediatr Adolesc Med*. 2002;156(5):460–7.
 25. Garcia JM, Gherpelli JLD, Leone CR. The role of spontaneous general movement assessment in the neurological outcome of cerebral lesions in preterm infants. *J Pediatr (Rio J)*. 2004;80(4):296–304.
 26. Geerdink, J. J., Hopkins B. Effects of Birthweight Status and Gestational Age on the Quality of General movements in preterm newborns. *Biol Neonate*. 1993;
 27. Geerdink, J.J., Hopkins B. Qualitative changes in general movements and their prognostic value in preterm infants. *Eur J Pediatr*. 1993;362–7.
 28. George JM, Fiori S, Fripp J, Pannek K, Guzzetta A, David M, et al. Relationship between very early brain structure and neuromotor, neurological and neurobehavioral function in infants born <31 weeks gestational age. *Early Hum Dev*.

- 2018;117(June):74–82.
29. Hadders-Algra M, Klip-van den Nieuwendijk AW, Martijn A, van Eyken LA. Assessment of general movements : towards a better evaluate brain function in young infants. *Dev Med Child Neurol*. 1997;39:89–99.
 30. Hadders-Algra M, Groothuis a M. Quality of general movements in infancy is related to neurological dysfunction, ADHD, and aggressive behaviour. *Dev Med Child Neurol* [Internet]. 1999;41(6):381–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10400172>
 31. Kakebeeke T, von Siebenthal K LR. Movement Quality in Preterm Infants Prior to Term. *Biol neonate*. 1998;145–54.
 32. Manacero SA, Marschik PB, Nunes ML, Einspieler C. Is it possible to predict the infant’s neurodevelopmental outcome at 14months of age by means of a single preterm assessment of General Movements? *Early Hum Dev* [Internet]. 2012;88(1):39–43. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0378378211002167>
 33. Mutlu A, Livanelioğlu A, Korkmaz A. Assessment of “general movements” in high-risk infants by Prechtl analysis during early intervention period in the first year of life. *Turk J Pediatr* [Internet]. 2010;52(6):630–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21428196>
 34. Nakajima Y, Einspieler C, Marschik PB, Bos AF, Prechtl HFR. Does a detailed assessment of poor repertoire general movements help to identify those infants who will develop normally? *Early Hum Dev* [Internet]. 2006;82(1):53–9. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0378378205001581>
 35. Olsen JE, Brown NC, Eeles AL, Lee KJ, Anderson PJ, Cheong JLY, et al. Trajectories of general movements from birth to term-equivalent age in infants born b 30 weeks ’ gestation. *Early Hum Dev*. 2015;91:683–8.
 36. Olsen JE, Brown NC, Eeles AL, Einspieler C, Lee KJ, Thompson DK, et al. Early general movements and brain magnetic resonance imaging at term-equivalent age in infants born < 30 weeks’ gestation. *Early Hum Dev* [Internet]. 2016;101:63–8. Available from: <http://dx.doi.org/10.1016/j.earlhumdev.2016.06.009>
 37. Olsen JE, Allinson LG, Doyle LW, Brown NC, Lee KJ, Eeles AL, et al. Preterm and term-equivalent age general movements and 1-year neurodevelopmental outcomes for infants born before 30 weeks’ gestation. *Dev Med Child Neurol*. 2018;60(1):47–53.
 38. Roescher M, Timmer A, Hitzert MM, de Vries NKS, Verhagen E a, Erwich JJHM, et al. Placental pathology and neurological morbidity in preterm infants during the first two weeks after birth. *Early Hum Dev* [Internet]. 2014;90(1):21–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24331826>
 39. Snider LM, Majnemer A, Mazer B, Campbell S, Bos AF. A comparison of the general movements assessment with traditional approaches to newborn and infant assessment: Concurrent validity. *Early Hum Dev* [Internet]. 2008;84(5):297–303. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0378378207001223>
 40. Spittle AJ, Walsh J, Olsen JE, McInnes E, Eeles AL, Brown NC, et al. Neurobehaviour and neurological development in the first month after birth for infants born between 32-42 weeks’ gestation. *Early Hum Dev* [Internet]. 2016;96:7–14. Available from: <http://dx.doi.org/10.1016/j.earlhumdev.2016.02.006>
 41. van Iersel P a M, Bakker SCM, Jonker AJH, Hadders-Algra M. Does perinatal

- asphyxia contribute to neurological dysfunction in preterm infants? *Early Hum Dev* [Internet]. 2010;86(7):457–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20605570>
42. Larroque B, Ancel PY, Marret S, Marchand L, André M, Arnaud C, et al. Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study. *Lancet*. 2008;371(9615):813–20.
 43. Chenouard A, Gascoin G, Gras-Le Guen C, Montcho Y, Roz?? JC, Flamant C. Neurodevelopmental impairment in preterm infants with late-onset infection: Not only in extremely preterm infants. *Eur J Pediatr*. 2014;173(8):1017–23.
 44. Martins PS, Mello RR de, Silva KS da. Bronchopulmonary dysplasia as a predictor factor for motor alteration at 6 months corrected age in premature infants. *Arq Neuropsiquiatr*. 2010;68(5):749–54.
 45. Butcher-Puech MC, Henderson-Smart DJ, Holley D, Lacey JL, Edwards DA. Relation between apnoea duration and type and neurological status of preterm infants. *Arch Dis Child*. 1985;60(10):953–8.
 46. Weisglas-Kuperus N, Baerts W, Fetter WPF, Sauer PJJ. Neonatal cerebral ultrasound, neonatal neurology and perinatal conditions as predictors of neurodevelopmental outcome in very low birthweight infants. *Early Hum Dev*. 1992;31(2):131–48.
 47. Petrini JR, Dias T, McCormick MC, Massolo ML, Green NS, Escobar GJ. Increased risk of adverse neurological development for late preterm infants. *J Pediatr* [Internet]. 2009;154(2):169–76. Available from: <http://www.sciencedirect.com/science/article/pii/S0022347608006999>
 48. Committee on Nutrition Pediatrics. Developmental Surveillance and Screening of Infants and Young Children. *Pediatrics*. 2001;106(2):346–9.
 49. Craciunoiu O, Holsti L. A Systematic Review of the Predictive Validity of Neurobehavioral Assessments During the Preterm Period. *Phys Occup Ther Pediatr* [Internet]. 2016;2638(June):1–16. Available from: <http://www.tandfonline.com/doi/full/10.1080/01942638.2016.1185501>
 50. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol* [Internet]. 2004;4(1):22. Available from: <http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-4-22>
 51. Chang SM, Matchar DB, Smetana GW, Umscheid C a, Gray R, Torchia MM. Methods Guide for Medical Test Reviews. *Methods Guid Med Test Rev* [Internet]. 2012;(12). Available from: http://www.effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/22834019%0Ahttp://www.ncbi.nlm.nih.gov/pubmed/22834019

Figure 1. Flow diagram depicting the study selection process

Table 1. Included Studies on Discriminative and Predictive Validity of the General Movements Assessment in Preterm Infants During the Preterm Period

Reference Evidence	N	Sampling criteria Participants characteristics	Brain abnormalities	Age of initial GMA	Outcome measure	Evaluated Relations	Accuracy measures (95% CI)
Albers and Jorch(15) Predictive	22	Infants of <32wk GA with AV GA: 25-31 ^a , BW: 490-1.740 ^a	IVH+PVL/hydrocephalus n=4	25–31	Neurological outcome based on the Münchner Funktionell at 1yr	Neurological outcome (Fisher's exact test, <i>p</i> =0.001)	Neurological outcome: sensitivity=89-100, specificity=33-83, PPV=56-82, NPV=90-100
Bos et al.(16) Discriminative	19	SGA infants GA: 31.4 (28–35) ^b , BW: 910 (595-1330) ^b	GMH1/2 ^c ; <i>n</i> =5/1, PVL1 ^d ; <i>n</i> =9	28–35		SGA (Fisher's exact test, <i>p</i> =0.001)	
Cioni et al.(17) Predictive	66	Infants with available assessments GA: 30.7 (26–36) ^e , BW: 1505 (850-2700) ^e	IVH1a2b/3 ^f ; <i>n</i> =6/4, PVL1/2/3-4 ^d ; <i>n</i> =23/6/17	26–36	Neurodevelopment based on Amiel-Tison and Grenier examination or Touwen's criteria and Griffiths scales at 2yr		Neurodevelopment: sensitivity=90.6, specificity=57.6, PPV=67.5, NPV=86.3
Cioni et al.(18) Discriminative	16	Brain-injured infants GA: 29.6 (26–35) ^e , BW: 1241 (700-2250) ^e	UIPE ^g ; <i>n</i> =16	26–35		UIPE (Fisher's exact test, <i>p</i> =0.005)	
Constantinou et al.(19) Predictive	130	VLBW infants of <32wk GA: 27.5 (23–32) ^e , BW: 922.0 (502-1460) ^e	PVL: <i>n</i> =11, IVH2-4: <i>n</i> =11	36	Amiel-Tison neurological assessment and BSID at 18mo		CP: sensitivity=62, specificity=69, PPV=29, NPV=90
de Vries et al.(20) Discriminative Predictive	19	ELBW infants GA: 27 (25–31.1) ^b , BW: 835 (520–995) ^b	GMH1/3 ^c ; <i>n</i> =2/1, TF: <i>n</i> =9, PVL1 ^d ; <i>n</i> =3	25.2-31.3	Paediatric neurodevelopmental follow-up and TINE at 20mo	GA: (<i>p</i> =0.02), SGA: (<i>p</i> =0.02), US (Fisher's exact test, <i>p</i> =0.033), AV: (<i>p</i> <0.05), sepsis: (<i>p</i> <0.005), hyperbilirubinemia: (<i>p</i> =0.035), neurodevelopment (<i>p</i> >0.05)	
de Vries and Bos(21) Discriminative Predictive	45	Infants of ≤32wk GA GA: 28.6 (25–32) ^b , BW: 865 (520-2035) ^b	IVH2 ^c ; <i>n</i> =2, PVL1 ^d : <i>n</i> =7	25–32	Paediatric examination and TINE at 24mo	GA: <i>r</i> _s =0.418 (<i>p</i> =0.004), BW: <i>r</i> _s =0.534 (<i>p</i> =0.00), SGA: (<i>p</i> =ns), BI (<i>p</i> =ns), neurological outcome: (<i>p</i> =ns)	Neurological outcome: sensitivity=80 (45–100), specificity=50 (32–68), PPV=21 (3–39), NPV=94(82–100)
Droit et al.(22) Discriminative Predictive	24	Infants with available assessments GA: 26–36 ^a , BW: 900-2785 ^a	TF: <i>n</i> =5, Hemorraghe1: <i>n</i> =1 hemorraghe2b+ PVL1: <i>n</i> =2, hemorrhage 2a+PVL2: <i>n</i> =1, hemorrhage3: <i>n</i> =1, PVL2/3 ^d ; <i>n</i> =2/6	31–35	Neurological outcome at 18mo		US: sensitivity=100, specificity=66.6, PPV=75, NPV=100; Neurological outcome: sensitivity=100, specificity=66.6, PPV=75, NPV=100
Ferrari et al.(23) Discriminative Predictive	29	Brain-injured infants GA: 26-36 ^a , BW: 880-2250 ^a	PF ^h ; <i>n</i> =7, GMH-IVH2-4 ^h ; <i>n</i> =6, GMH-IVH and/or PVL ^h ; <i>n</i> =16	26–36	Neurological outcome based on Amiel-Tison neurological examination, Touwen's criteria, Griffiths or Brunet Lezine and Terman Merill scales at 24mo		BI: sensitivity=100, specificity=92.8, PPV=96, NPV=100; Neurological outcome: sensitivity=100, specificity=59, PPV=70, NPV=100
Ferrari et al.(24) Predictive	84	Brain-injured infants AG: 30.2 (2.7) ⁱ , BW: 1410.1 (456.7) ⁱ	White matter abnormal: <i>n</i> =68, IVH3/3+ ^j ; <i>n</i> =16	<37	Neurodevelopmental outcome according to the Griffiths scales at 2-3yr		CP: sensitivity=100, specificity=38, PPV=63, NPV=100
Garcia et al.(25) Discriminative Predictive	40	Infants of <35wk GA AG: 31 (26–34) ^e , BW: 1100 (630-1530) ^e	PVH1/2/3 ^c : <i>n</i> =1/2/6, PVL: <i>n</i> =4, ventri-, periventricular abnormality: <i>n</i> =3	25	Neurological outcome based on neurological examination and DDST at 1yr	US (McNemar's test, <i>p</i> =0.41), neurological outcome (McNemar's test, <i>p</i> =0.09)	US: sensitivity=71, specificity=42, PPV=25, NPV=85; Neurological outcome: sensitivity=100, specificity=44, PPV=36,

Geerdink and Hopkins(26) Discriminative	37	Infants without medical complications SGA group $n=15$, GA: 31.6 (29-34) ^b , AGA group $n=22$, GA: 32.7 (27-34) ^b	PE: $n=1$, IVH1/2: $n=1/2$	35		GA: $\chi^2=8.22$ ($p<0.02$), SGA/AGA: $\chi^2=10.03$, ($p<0.01$), obstetrical optimality scores: $F=5.31$ ($p=0.01$)	NPV=100
Geerdink and Hopkins(27) Predictive	35	Infants without medical complications GA: 32 (27-34) ^b	PE: $n=1$, IVH1/2: $n=1/2$	35	Neurological outcome according to TINE and BSID at 1yr		Neurological outcome: sensitivity=40-60, specificity=58, PPV=29-100, NPV=47
George et al(28) Discriminative	119	Infants of <31wk GA GA: 28 (26-29) ^b ; BW: 1093 (494-1886) ^b	IVH: $n=30$, PVL: $n=4$: Hydrocephalus+ IVH3/4: $n=4$	30-32	Structural MRI at 30-32wk (early) and 40-42wk (term)	Structural MRI score (Regression analysis, $p=ns$). Early MRI of the cerebellum ($p<0.001$) US (Fisher's exact test, $p<0.05$),	
Hadders-Algra et al.(29) Discriminative	16	Infants with abnormal GM GA: 26-36 ^a , BW: 850-2750 ^a	Hydrocephalus: $n=1$, cerebral oedema: $n=1$, PVL1/2/3 ^d : $n=5/1/1$, PVH2 ^e : $n=1$	<38			
Hadders-Algra and Groothuis(30) Predictive	52	Infants with abnormal GM Preterm infants $n=10$, GA: 30 (26-36) ^b , BW: 1438 (548) ^f	PVL1/2/3 ^d : $n=10/2/1$, posthydrocephalus state: $n=1$	<38	Neurological examination based on Hempel assessment and TINE, behavioural outcome by means of parental questionnaires at 4yr-9yr	Neurological development (Fisher's exact test, $p>0.05$), behavioural outcome: (Fisher's exact test, $p>0.05$)	
Kakebeeke et al.(31) Discriminative	56	Infants of <36wk GA high-risk-group: $n=18$, GA: 28.6 (26.1-31.3) ^f , BW: 1160 (730-1710) ^f , low-risk-group $n=18$, GA: 34.5 (33.3-35.9) ^f , BW: 2087 (1520-2800) ^f	High-risk group: IVH1/2/3 ^d : $n=2/3/1$, PVL1 ^d : $n=3$, PVL1 ^d +IVH1 ^e : $n=5$, PVL2 ^d +IVH3 ^e : $n=1$, low-risk group: PVL1 ^d : $n=2$	26.1-35.9		GA (F , $p<0.05$), US (F , $p<0.05$), US: $r_s=0.51$, -0.52, -0.53, -0.47 ($p<0.05$), US: $r_s=-0.46$, -0.41 ($p>0.05$), low-risk-group/control-group (H , $p<0.001$), high-risk-group/low-risk-group (H , $p<0.001$)	
Manacero et al.(32) Discriminative Predictive	37	Infants of ≤ 33 wk GA GA: 31 (24-33) ^b , BW: 1455 (640-2050) ^b	IVH1 ^b : $n=5$, PVL3 ^d : $n=1$	34	Developmental status according to the AIMS and the PEDI at 14mo	GA (H , $p=ns$), BW (H , $p=ns$), US ($p=ns$), VA days ($p=ns$); AIMS-centile: (H , $p<0.05$), $r_s=0.43$ ($p=0.01$) AIMS-categories: $V=0.58$ ($p<0.01$), PEDI-motility: $r_s=0.41$ ($p<0.05$), PEDI-selfcare: $r_s=0.22$ ($p=ns$), PEDI-social: $r_s=0.30$ ($p=ns$) US: Kappa (PABAK)=0,86	
Mutlu et al.(33) Discriminative	28	LBW infants GA: 30.9 (2.6) ^j , BW: 1488 (420) ^j	IVH/PVL: $n=2$	34.8			
Nakajima et al.(34) Predictive	18	Infants with abnormal GM GA: 28 (25-33) ^b , BW: 942 (595-1550) ^b	PF: $n=12$, PF+IVH1: $n=3$ IVH1: $n=1$, PVL2 ^d : $n=1$	27-33	Neurological outcome according to TINE at 8-10yr	Neurological outcome (H , $p=ns$)	
Olsen et al.(35) Discriminative	145	Infants of ≤ 30 wk GA GA: 27.7 (1.5) ^j , BW: 1024 (263) ^j	IVH1-2/3-4: $n=24/4$, PVL: $n=1$	<32		GA ($p<0.001$), BWz score ($p=ns$), BPD ($p=0.001$), Sepsis/NEC: (Regression analysis, $p<0.001$)	
Olsen et al.(36) Discriminative	149	Infants of ≤ 30 wk GA GA: 28 (1.3) ^j , BW: 1080 (247) ^j	IVH3-4: $n=3$, PVL+IVH3-4: $n=1$	<32	Brain abnormality and growth in brain MRI at 38-44wk	Cortical grey matter abnormality (Regression analysis, $p<0.03$), deep grey matter abnormality (Regression analysis, $p=0.02$), increased interhemispheric distance (Regression analysis, $p<0.02$)	
Olsen et al(37) Predictive	137	Infants of <30wk GA GA: 27.8 (1.5) ^j , BW 1031 (262) ^j	IVH3/4: $n=3$	VPT=<32 MPT=32-33 LPT=34-36	Neurodevelopmental outcomes based on the AIMS, the NSMDA and TINE at 12mo	Total scores of the AIMS, the NSMDA and TINE (Regression analysis, $p<0.05$)	AIMS: sensitivity=89(77-95), 81(69-90), 85(69-94), specificity=19(10-31), 26(16-41), 8(2-23), PPV=53(43-62), 55(44-66), 50(38-62), NPV=61(36-82), 56(35-75), 33(9-69); NSMDA: sensitivity=90(82-95), 82(72-89), 90(78-96);

specificity=32(17–52), 36(19–56), 16(4–40); PPV=81(72–88), 79(69–87), 76(64–86); NPV=50(27–73), 40(22–61), 33(9–69); TINE: sensitivity=90(80–95), 86(74–93), 93(81–98); specificity=23(13–28), 32(20–47), 17(7–36), PPV=63(53–73), 63(52–73), 64(51–75), NPV=61(36–82), 63(41–80), 63(26–90)

Roescher et al. (38) Discriminative	52	Infants of <32wk GA GA: 29.1 (25.1–31.7) ^b , BW: 1180 (560–2250) ^b	Hemorrhage1–2: n=7, hemorrhage3–4: n=4	<32		GA (<i>U</i> , <i>p</i> <0.01; rs, <i>p</i> <0.01), BW (rs, <i>p</i> <0.01), SGA (<i>p</i> >0.05), AGA (<i>p</i> >0.05), AV (χ^2 for trends <i>p</i> <0.01), BI: (<i>p</i> =ns), placental pathology: (<i>p</i> =ns) AV days (Regression analysis, <i>p</i> =0.06)
Snider et al.(39) Discriminative	100	Infants of <32wk GA GA: 28.6 (2.5) ⁱ , BW: 1075 (258.4) ⁱ	IVH1/2/3/4: n=7/1/2/1	34		GA in MPT group (Regression analysis, <i>p</i> <0.05, OR=0.20 (0.06–0.67))
Spittle et al.(40) Discriminative	410	MPT/LPT infants n=80/129, GA: 33 (0.6) ^j /35.2 (0.8) ^j , BW: 1887 (364) ^j /2329 (450) ^j		34.4/36.4		GA in MPT group (Regression analysis, <i>p</i> <0.05, OR=0.20 (0.06–0.67))
van Iersel et al.(41) Discriminative Predictive	17	Infants with perinatal asphyxia GA: 32 (28–34) ^b , BW: 1817 (611) ⁱ	PVL1/2/3/4 ^d with/without GMH- IVH ^c : n=6/3/1/0	<38	CP according paediatric assessment and TINE at 2yr.	GA (Fisher's exact test, <i>p</i> =0.04; Regression analysis, <i>p</i> =0.1), OR=0.16 (0.02–1.49), perinatal asphyxia (<i>U</i> , <i>p</i> >0.05), asphyxied-group/non-asphyxied group (<i>p</i> >0.05), respiratory problems (Fisher's exact test, <i>p</i> =0.02, Regression analysis, <i>p</i> =0.02, OR=6.29 (1.42–27.76), CP (χ^2 for trends <i>p</i> >0.05)

Note: N; sample, CI, confidence interval is a values range considered to be plausible for the population; AV, artificial ventilation; ^arange; IVH, intraventricular hemorrhage; PVL, periventricular leucomalacia; Fisher's exact test is used to measure relation between nominal cross-tabulated variables; sensitivity (%), proportion of true-positives correctly identified; specificity (%), proportion of true-negatives correctly identified; PPV, Positive Predictive Value; NPV, Negative Predictive Value; SGA, small for gestational age; ^bmedian(range); GMH, germinal matrix hemorrhage; ^cclassification according to Papile et al. (1978); ^dclassification according to de Vries et al. (1992); ^emean(range); ^fclassification according to Levene et al. (1982); UIPE, unilateral intraparenchymal echodensity; ^gDefinition according to Guzzetta et al. (1986); VLBW, very low birth weight; BSID, Bayley Scales of Infant Development; CP, cerebral palsy; ELBW, extremely low birth weight; TF, transient flaring; TINE, Touwen Infant Neurological Evaluation; rs, Spearman coefficient indicates the degree of relation between ordinal variables by a range from -1 to 1; US, ultrasonography; BI, brain injury; PF, prolonged flares; ^hclassification according to Trounce et al. (1988); ⁱmean(SD), ^jclassification according to Volpe (1989); PVH, periventricular hemorrhage; DDST, Denver Developmental Screening Test; McNemar's test evaluates the relation between ordinal variables in dependent samples; AGA, appropriate for gestational age; PE, periventricular echodensity; χ^2 , the Chi-square test measures the independence between categorical cross-tabulated variables; *F*, report of variance analysis (ANOVA); MRI, Magnetic resonance imaging; Regression analysis, it's a set of methods for estimating relation among a dependent variable and one or more independent variable. *H*, the Kruskal-Wallis is a nonparametric test measuring differences between samples; ^kclassification according to Volpe (2008); AIMS, Alberta Infant Motor Scales; PEDI, Paediatric Evaluation of Disability Inventory; *V*, the Cramer's test measures, with a range from 0 to 1, the degree of relation between nominal variables; LBW, low birth weight; Kappa (PABAK), the Kappa test measures, with a range from 0 to 1, the agreement of conditions with reference standard; BDP, bronchopulmonary dysplasia; NEC, necrotizing enterocolitis; VPT, very preterm; MPT, moderate preterm; LPT, late preterm; NSMDA, Neurological, Sensory, Motor, Developmental Assessment; χ^2 for trends tests association between a nominal variable with two levels and an ordinal variable; OR, odds ratio indicates the probability that an outcome does occur divided by the probability that the outcome does not occur; *U*, Mann-Whitney test evaluates mean differences for ordinal variables.

Table 2. Quality of included studies based on ‘The Critical Review form Quantitative Studies’

Study	Purpose of study clearly stated	Relevant literature review	Study design appropriate	No biases present	Sample described in detail	Reliability of assessments tool should be reported by the authors	Validity of assessments tool should be reported by the authors	Results reported in terms of statistical significance	Analysis of results was appropriate	Clinical importance of results were reported	Loss data were reported	Conclusions were appropriate	Score	Issue
Albers S, Jorch G, 1994	1	1	1	0	1	0	0	1	1	1	1	1	9	Specifying sampling method Reliability testing Reporting measures for population estimate
Bos et al., 1997	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis Sampling based on assessment available Reporting measures for population estimate
Cioni et al.,1997	1	1	1	0	1	0	0	1	1	1	1	1	9	Sampling based on assessment available Reporting measures for population estimate
Cioni et al., 2000	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Sampling based on assessment available Reporting measures for population estimate
Constantinou et al., 2007	1	1	1	0	1	0	0	1	1	1	1	1	9	Describing assessment independence Reliability testing Reporting measures for population estimate
de Vries et al., 2008	1	1	1	0	1	0	0	1	1	1	1	1	9	Specifying sampling method Describing assessment independence Reliability testing Reporting measures for population estimate
de Vries N, and Bos A, 2010	1	1	1	0	1	0,5	0	1	1	1	1	1	9,5	Formulating hypothesis. Specifying sampling method Reliability testing Reporting measures for population estimate
Droit et al.,1996	1	1	1	0	1	0	0	0,5	1	1	1	1	8,5	Formulating hypothesis. Sampling based on assessment available Describing assessment Reporting measures for population estimate
Ferrari F, et al.,1990	1	1	1	0	0,5	0	0	1	1	1	1	1	8,5	Formulating hypothesis Specifying sampling method Describing assessment independence Reliability testing Reporting measures for population estimate
Ferrari et al., 2002	1	1	1	0	1	1	0	1	1	1	1	1	10	Formulating hypothesis Sampling based on assessment available

Garcia et al., 2004	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Reliability testing Significant loss of data Reporting measures for population estimate
Geerdink J, Hopkins, B. 1993	1	1	1	0	1	0	0	1	1	1	1	1	9	Specifying sampling method Describing assessment independence? Reporting measures for population estimate
Geerdink J, Hopkins, B. 1993	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Specifying sampling method Reporting measures for population estimate
George et al (2018)	1	1	1	1	1	0,5	0,5	1	1	1	1	1	11	Reliability testing
Hadders-Algra et al., 1997	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Sampling based on assessment available Reporting measures for population estimate
Hadders-Algra M, Groothuis A, 1999	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Sampling based on assessment available Reporting measures for population estimate
Kakebeek e T, von Siebenthal K, 1998	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis Reporting measures for population estimate
Manacero et al., 2012	1	1	1	0	1	0	1	1	1	1	1	1	10	Formulating hypothesis. Describing assessment independence Reporting measures for population estimate
Mutlu et al., 2010	1	1	1	0	1	0	1	1	1	1	1	1	10	Formulating hypothesis Specifying sampling method Significant loss of data ? Reporting measures for population estimate
Nakajima et al 2006	1	1	1	0	1	0	1	1	1	1	1	1	10	Formulating hypothesis. Sampling based on assessment available Reporting measures for population estimate
Olsen et al., 2015	1	1	1	1	1	0	0	1	1	1	1	1	10	Specifying sampling method
Olsen et al., 2016	1	1	1	0	1	0	0	1	1	1	1	1	9	Specifying sampling method Significant loss of data ?
Olsen et al (2018)	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis
Roescher et al., 2014	1	1	1	0	1	1	0	1	1	1	1	1	10	Specifying sampling method Reliability testing Reporting measures for population estimate
Snider et al., 2008	1	1	1	1	1	0	1	1	1	1	1	1	11	

Spittle et al., 2016	1	1	1	0	1	0	0	1	1	1	1	1	9	Specifying sampling method Assessment independence was unclear
van Iersel et al., 2010	1	1	1	0	1	0	0	1	1	1	1	1	9	Formulating hypothesis. Specifying sampling method

Note: Modified by: Burger M, Louw Q a. The predictive validity of general movements – A systematic review. Eur J Paediatr Neurol. Elsevier Ltd; 2009; 13(5):408–20. Original Version in: Law, M. Stewart, D. Pollock, N. Letts, L. Bosch, J & Westmoreland M. Guidelines for Critical Review form Quantitative Studies. Quant Rev Form-Guidelines [Accessed in March 2018]. 1998; 1–11. Available from: <http://srs.mcmaster.ca/research/evidence-based-practice-research-group/>; very high (>80%), high (>60%-80%<), and low (<60%). Criteria on assessment of the risk of bias are detailed in the text

Annexe 4

Cette dernière partie de notre travail, réalisée en 2019, est le résultat d'une collaboration avec l'ISIR (Institut des Systèmes Intelligents et de Robotique, Unité Mixte de Recherche commune à Sorbonne Université et au Centre National de la Recherche Scientifique (CNRS)). La synthèse présentée ici est le travail de Loïc Goasguen qui a effectué son stage de master sous la direction de Catherine Achard.

Rapport de stage

Analyse vidéographique des mouvements généraux des nouveau-nés prématurés à haut risque

par Loïc Goasguen
Master 1



Tutrice de stage : Catherine Achard
Tuteur académique : Bruno Gas

17 Mai - 22 Juillet 2019

Remerciements

Je tiens à remercier les personnes qui ont contribué à la réussite de mon stage.

En premier lieu, je souhaite remercier Madame Catherine Achard, ma tutrice de stage, pour son accueil, la confiance qu'elle m'a accordée, son investissement dans mon projet de stage ainsi que pour sa relecture de mon rapport.

Je remercie également l'ISIR et plus particulièrement l'équipe PIRoS pour leur accueil.

Enfin j'en profite également pour remercier l'équipe pédagogique du Master Sciences Pour l'Ingénieur de Sorbonne Université.

Table des matières

1	Introduction	1
2	Présentation du laboratoire et de l'équipe de recherche	2
2.1	L'ISIR.....	2
2.2	L'équipe PIROS.....	2
2.2.1	Perception de et chez l'humain.....	3
2.2.2	Interaction Sociale.....	3
2.2.3	Psychopathologie.....	3
3	Présentation du sujet	4
3.1	Contexte du stage	4
3.2	Contexte médical.....	4
3.2.1	Les mouvements généraux spontanés ou GMs	4
3.2.2	Évaluation de la qualité des GMs	5
3.3	Problématiques et enjeux	5
3.4	Objectifs.....	5
3.5	Méthodologie	5
4	L'intelligence artificielle et l'apprentissage profond	7
4.1	L'apprentissage machine.....	7
4.2	Le perceptron.....	8
4.3	Le perceptron multicouches	9
5	État de l'art des méthodes d'apprentissage profond	11
5.1	Les réseaux de neurones convolutifs.....	11
5.1.1	Les réseaux de neurones spatiaux convolutifs . .	11
5.1.2	Les réseaux de neurones temporels convolutifs 3D .	12
5.1.3	Le modèle C3D	13
5.2	Les réseaux de neurones récurrents.....	13
5.2.1	Les RNNs.....	13
5.2.2	Les LSTMs	15
6	Exposé des travaux	17
6.1	1 ^{er} Modèle : Analyse des caractéristiques spatio-temporelles via une approche double utilisant un modèle C3D et un LSTM.....	17
6.1.1	Modèle.....	18
6.1.2	Implémentation du modèle.....	19
6.1.3	Réduction de dimension et traitements additionnels .	23
6.1.4	Conclusion	24
6.2	Encodeur TCN	25
6.2.1	Modèle.....	25
6.2.2	Implémentation.....	26
6.2.3	Conclusion	26
6.3	Estimation du mouvement via flot optique et différence d'images successives	27
6.3.1	Extraction des informations de mouvements.....	27
6.3.2	C3D et LSTM.....	28
6.3.3	Création de descripteurs et LSTM	29
6.3.4	Conclusion	30

6.4	Résultats obtenus et comparaison avec les objectifs	31
6.5	Difficultés rencontrées	31
7	Conclusion	32

1 Introduction

Il a été montré que les mouvements généraux spontanés sont révélateurs de l'avenir neuro-développemental des nouveaux nés. Des recherches ont été effectuées afin de mettre au point une technique d'analyse capable de caractériser leur qualité et ainsi d'utiliser ces informations pour accompagner les enfants à risque. Cette technique est très efficace lorsqu'elle est effectuée par des humains mais elle est complexe et nécessite d'être formés. Dans ce projet de stage on souhaite développer un outil informatique utilisant des algorithmes d'apprentissage profond afin de caractériser des séquences vidéos de mouvement généraux sur des enfants prématurés à haut-risque.

C'est dans ce domaine que j'ai effectué mon stage de Master 1 sur une période de 2 mois allant de Mai à Juillet 2019 au sein de l'équipe PIRoS de l'ISIR. J'étais encadré par Catherine Achard en tant que responsable pour le laboratoire et par Bruno Gas en tant que responsable académique. Pendant ce stage j'ai pu découvrir les méthodes d'apprentissage profond (Deep Learning) via différentes techniques appliquées au traitement d'image et de vidéos. J'ai pu étudier différents modèles et les mettre en place afin de trouver un moyen de répondre au problème donné.

Premièrement, nous présenterons le laboratoire et l'équipe que j'ai intégrée. Ensuite nous développerons le sujet et les objectifs qu'il implique en détaillant le contexte médical dans lequel cet outil s'intègre. Puis nous effectuerons des rappels ainsi qu'un état de l'art sur les méthodes d'apprentissage profond appliquées au traitement d'image et de vidéos qui seront utilisées par la suite. Enfin nous présenterons le détail du travail effectué pendant ce stage en développant les 3 modèles utilisés avant de conclure ce rapport.

2 Présentation du laboratoire et de l'équipe de recherche

2.1 L'ISIR

L'Institut des Systèmes Intelligents et de Robotique est un laboratoire de recherche pluridisciplinaire qui rassemble des chercheurs et enseignants-chercheurs relevant de différentes disciplines des Sciences de l'Ingénieur et de l'Information ainsi que des Sciences du Vivant. L'ISIR est une Unité Mixte de Recherche commune à Sorbonne Université et au Centre National de la Recherche Scientifique (CNRS). Il est rattaché d'une part à la faculté d'Ingénierie de Sorbonne Université et d'autre part à l'Institut des Sciences de l'Information et de leurs Interactions (INS2I) du CNRS. L'Institut national de la santé et de la recherche médicale (INSERM) est également tutelle de l'une de ses équipes.

Les thématiques de recherche abordées par l'ISIR sont centrées sur la robotique et les interactions humain-machine. Ainsi les problématiques couvrent les interactions cognitives et physiques à différentes échelles, la perception multimodale, la manipulation, la robotique autonome avec les problématiques de navigation ou de prise de décision, l'optimisation et l'apprentissage ainsi que les architectures cognitives. Le laboratoire se découpe en 5 équipes : AGATHE, AMAC, SYROCO, PIRoS et Interaction multi-échelles.

L'équipe Assistance aux Gestes et Applications THÉrapeutiques (AGATHE) s'intéresse à la conception et à la commande de dispositifs robotiques pour l'assistance aux gestes et mouvements humains. Ces recherches se modélisent sous la forme d'assistance aux interventions médico-chirurgicales ainsi que d'assistance aux personnes souffrant d'un déficit moteur. L'équipe Architectures et Modèles pour l'Adaptation de la Cognition (AMAC) s'intéresse à l'élaboration de modèles des fonctions perceptives, cognitives et motrices ainsi qu'à la synthèse d'architectures de contrôle dans une perspective interactive. L'équipe SYstèmes RObotiques COmplexes (SYROCO) travaille sur le développement de méthodes de design et de contrôle pour la robotique mobile et les systèmes de manipulation. L'équipe Interactions multi-échelles s'intéresse aux interactions entre l'environnement physique et numérique. Plusieurs approches sont abordées comme l'haptique, la visualisation, l'audio ainsi que la réalité virtuelle et augmentée.

Enfin l'équipe Perception Interaction et Robotique Sociale est l'équipe que j'ai intégrée pour effectuer mon stage. Dans la prochaine partie je vais tâcher de décrire plus profondément ses activités.

2.2 L'équipe PIRoS

L'équipe PIRoS (Perception Interaction et Robotique Sociale) est portée par Mohamed Chetouani. Ses recherches dans le domaine de la Perception sociale, de l'Interaction sociale et de la Robotique sociale reposent sur des méthodologies de l'apprentissage statistique, du traitement du signal et de la vision par ordinateur pour une part, et de la psychologie, des neurosciences, des sciences cognitives et de la psychiatrie d'autre part. Cette composition pluridisciplinaire permet à l'équipe d'aborder des dimensions importantes de la chaîne de l'interaction humain-humain et humain-machine : de la perception à la génération de comportements en passant par la modélisation interpersonnelle ou encore la collecte de données dans un contexte social.

Les thématiques abordées par l'équipe PIROS peuvent se résumer en trois catégories : la perception de et chez l'humain, l'interaction sociale et les psychopathologies.

2.2.1 Perception de et chez l'humain

L'équipe s'intéresse à deux aspects de la perception. Premièrement comprendre comment l'humain perçoit et interprète les signaux sociaux multimodaux (visage, voix, geste, toucher) de son environnement. En utilisant des méthodes de psychologie expérimentale ainsi que des approches computationnelles (méthodes d'apprentissage automatique) l'équipe modélise les différentes stratégies perceptives. Inversement, du point de vue de la perception de l'humain par la machine, ils proposent des méthodes d'analyse automatique du comportement humain : expressions faciales, gestes, voix, activité neurophysiologique. Ces analyses reposent sur la conception de méthodes d'apprentissage statistique, de traitement du signal et de vision par ordinateur pour le traitement de signaux mono et multimodaux.

2.2.2 Interaction Sociale

Cette thématique étudie l'interaction entre un être humain et un agent social artificiel. Elle cherche à modéliser des signaux sociaux-émotionnels à travers différentes modalités (e.g. voix, regard, expressions faciales, toucher social). Les thématiques de recherche portent sur l'imitation motrice avec des robots humanoïdes, le regard social capté par eye-tracking, la synchronisation entre partenaires humains et agents en lien avec leur niveau d'engagement, l'action collaborative médiée par des interfaces haptiques ou l'apprentissage interactif. L'équipe utilise des méthodes issues de la psychologie expérimentale ainsi que des approches computationnelles et en particulier l'apprentissage statistique.

2.2.3 Psychopathologie

Enfin l'équipe PIROS abordent aussi l'étude de la psychopathologie au moyen de l'Interaction Sociale Humain-Machine sur plusieurs plans. D'abord l'imagerie comportementale afin d'améliorer la compréhension des mécanismes psychopathologiques. Ensuite une approche dimensionnelle de la psychopathologie. Les chercheurs étudient les principales dimensions du développement grâce aux méthodes ISHM. L'équipe travaille aussi sur la robotique sociale développementale qui permet par exemple de définir une signature sociale de l'autisme. Enfin ils développent aussi des applications cliniques comme des outils de diagnostic ou des méthodes de traitement thérapeutique.

C'est dans le contexte de développement d'un outil de diagnostic que le projet de ce stage se place.

3 Présentation du sujet

3.1 Contexte du stage

Dans le cadre du projet de recherche intitulé "Analyse vidéographique de la motricité spontanée" le Pôle de périnatalité de l'Hôpital Armand Trousseau, le laboratoire CHArt de l'École Pratique des Hautes Études et le service de pédiatrie et réanimation néonatales de l'Hôpital Universitaire Necker-Enfants Malades ont conduit une étude observationnelle.

Cette étude se base sur des enregistrements vidéo de la motricité spontanée et de la posture des nouveau-nés. Le but de cette étude est d'utiliser les enregistrements vidéo récupérés afin de développer un outil informatique capable d'identifier automatiquement les anomalies motrices précoces chez le nouveau-né prématuré à haut risque. Cet algorithme doit faire appel à des algorithmes informatiques et de traitement d'images développés à partir de l'information sur la motricité spontanée de l'échantillon constitué par les nouveau-nés participants à l'étude. Pour effectuer ce travail il est demandé à l'ISIR le développement d'un système d'apprentissage supervisé capable de différencier des séquences de mouvements normaux et anormaux.

Ce travail est le sujet de mon stage. Je vais poursuivre en décrivant plus précisément le travail qui m'a été demandé en évoquant le contexte médical du sujet et décrivant l'objectif à atteindre.

3.2 Contexte médical

3.2.1 Les mouvements généraux spontanés ou GMs

Les progrès récents des soins apportés pendant la grossesse et aux nouveaux nés couplés à la baisse de la mortalité périnatale ont conduit à augmenter le nombre d'enfants à haut risque de trouble du développement. Cependant il est difficile de prévoir dès le très jeune âge quel enfant va développer une infirmité motrice cérébrale (IMC), une « maladresse », des difficultés d'apprentissage scolaire ou un déficit d'attention avec hyperactivité (ADHD). Un dysfonctionnement peut disparaître pendant l'enfance comme apparaître spontanément augmentant les difficultés du pronostic, ceci est dû aux caractéristiques de développement du système nerveux. La nécessité d'innover dans les méthodes est venue de l'hétérogénéité des résultats obtenus par les techniques utilisées, d'où le besoin de trouver des méthodes plus poussées et précises dans la description des comportements de l'enfant [1].

Une technique d'évaluation de la qualité des mouvements généraux spontanés (GMs) a été développée pour évaluer l'intégrité du fonctionnement cérébral du nouveau-né et du jeune nourrisson [1]. Cette technique initiée par Prechtl en 1990 [2] repose sur l'évaluation qualitative de la motricité globale spontanée. Il a découvert que la qualité de la motilité spontanée, spécifiquement la qualité du « mouvement global » donne une vision précise de l'état fonctionnel du système nerveux et du fœtus du nouveau-né [2].

3.2.2 Évaluation de la qualité des GMs

Suite à ces découvertes est venu le besoin de créer une méthode d'évaluation de la qualité des GMs. Cette méthode est introduite dans [1]. Hadder-Alga commence par décrire les GMs normaux et comment ils sont répartis en trois catégories : pré-terme (avant 36 à 38 sem. d'âge post-menstruel), writhing (de 36 à 38 sem. d'âge post-menstruel) et figety (de 6 à 8 sem. après le terme). Chacune de ces périodes contient des caractéristiques qui lui sont propres. Afin d'évaluer la qualité des GMs, 3 critères sont étudiés sur les mouvements de l'enfant pendant une certaine période : la complexité ou variation spatiale, la variation ou variation temporelle et la fluidité. Dans [1] est ensuite explicité comment ces paramètres permettent d'évaluer la qualité des GMs. On en retient qu'il est donc possible d'associer à une séquence vidéo d'un enfant une évaluation sur l'"anormalité" des mouvements effectués.

3.3 Problématiques et enjeux

De ces documents ressort une problématique, s'il est possible d'évaluer la qualité des GMs il est tout de même nécessaire de former du personnel à une tâche qui est complexe et chronophage. On peut en conclure qu'il serait bénéfique de développer un outil d'analyse automatique de vidéos qui évaluerait l'anormalité des mouvements d'un enfant.

Dans le cadre de la recherche observationnelle : "Analyse vidéographique de la motricité spontanée" promu par l'École Pratique des Hautes études il est souhaité le développement d'un outil d'analyse supervisé. Ce programme permettrait de rendre accessible les tests à un plus grand nombre de personnes et aussi de fournir plus d'outils au personnel médical afin de diagnostiquer et d'accompagner leurs patients.

3.4 Objectifs

Pour cela, l'objectif est de développer un outil informatique qui fait appel à des algorithmes d'intelligence artificielle pour analyser la posture et l'activité motrice globale spontanée du nourrisson, l'attitude des bras et jambes, ainsi que la position des mains et des pieds. Le développement d'un tel algorithme a un objectif double. D'une part il permettra de rendre accessible l'évaluation des GMs mais il permettra aussi de vérifier expérimentalement la validité de cette technique. Pour cela, dans un premier temps le but est d'effectuer une classification binaire qui répertorie si une séquence de mouvement est normale ou anormale et dans un deuxième temps de donner un score d'anormalité à la séquence.

3.5 Méthodologie

Afin d'entraîner notre modèle, l'École Pratique des Haute Études nous a fourni un ensemble de données composé de vidéos de 82 enfants prématurés nés entre 28 et 37 semaines d'aménorrhée¹.

Des enfants ont été filmés pendant 10 minutes lors de la période pré terme (A), à l'âge du terme (B) et entre 9 et 15 semaines après terme (C). 820 séquences motrices ont été extraites de ces vidéos dont un exemple est illustré Figure 1. À partir de ces vidéos

1. L'aménorrhée caractérise l'absence de règles chez la femme en âge de procréer. L'aménorrhée se rencontre en situation normale pendant la grossesse mais peut aussi se produire de manière pathologique au cours de la période de fécondité, signe le plus souvent d'une maladie sous-jacente.

deux experts ont analysé la complexité, la variation et la fluidité des GMs en vue de les classer en normaux ou anormaux.



Figure 1 - Extraits GMs. Passages des vidéos fournies montrant des séquences de mouvements généraux spontanés chez des nourrissons.

Ces données seront fournies aux algorithmes d'intelligence artificielle qui tâcheront de généraliser les caractéristiques qui font qu'une séquence de mouvement est plus ou moins anormale. Dans l'ensemble des vidéos fournies, les vidéos classées comme anormales entre dans la catégorie "pauvre répertoire" où les séquences de mouvements sont répétitives. Elles ne présentent ni la variabilité ni la complexité observée dans les GMs normaux. La vitesse de ce type de mouvement est également monotone et la fluidité est diminuée. Ces mouvements ont en particulier été observés chez les enfants ayant des lésions cérébrales détectées par échographie transfontanellaire². Cependant, la valeur prédictive du devenir neurologique de ces enfants est plutôt faible [3, 4].

Ces données impliquent donc que notre modèle ne pourra prédire que des séquences vidéo où l'enfant manifeste des GMS anormaux de type "pauvre répertoire". Avec un ensemble contenant un plus grand nombre d'exemple des autres catégories (Chaotic, Cramed-synchronised, Abnormal fidgety, Absent fidgety) il serait possible d'associer chaque séquence anormale à sa cause, voire de classifier directement selon les diverses catégories.

2. L'échographie transfontanellaire est un examen d'imagerie médicale réalisé chez le nourrisson. Cette technique non invasive permet de visualiser le cerveau de l'enfant et de vérifier l'absence d'anomalies.

4 L'intelligence artificielle et l'apprentissage profond

Afin de développer un outil automatique d'analyse de vidéos à partir d'algorithmes d'apprentissage profond, il est nécessaire de rappeler ce qu'est l'apprentissage profond et quels outils apporte cette technique pour le traitement de vidéos. Pour commencer nous introduisons quelques éléments d'apprentissage machine.

4.1 L'apprentissage machine

Dans l'aire du big data toutes les informations sont présentes dans des quantités plus grandes que jamais. Ces quantités de données nécessitent des méthodes automatiques d'analyses, c'est ce que l'apprentissage machine permet de résoudre. Plus généralement on peut parler de l'apprentissage machine comme de l'ensemble des méthodes capables de détecter automatiquement les motifs dans les données. Ces motifs, une fois détectés, peuvent permettre de prédire les futures données ou effectuer diverses tâches de prise de décision sous incertitude [5].

Il existe trois principales approches d'apprentissage machine. L'apprentissage supervisé a pour but de faire correspondre des sorties y à des entrées x grâce à des ensembles d'entrées/sorties annotées. On divise souvent le problème en deux (Figure 2) : lorsque les réponses peuvent être divisées en catégories il s'agit d'un problème de classification. Quand la sortie y est une valeur réelle, il s'agit d'un problème de régression. La seconde approche est l'apprentissage non supervisé. Ici on ne reçoit que des entrées et le but est que l'algorithme trouve de lui-même des structures sous-jacentes dans les données. Le problème est moins bien défini car on ne sait pas ce que l'on recherche [5]. Enfin il existe une troisième forme appelée apprentissage par renforcement qui repose sur un système de récompenses pour déterminer ce qui doit être appris [5].

À l'intérieur de ces catégories, il existe une multitude de méthodes d'apprentissage machine mais la méthode la plus populaire depuis quelques années tourne autour des réseaux de neurones profond plus connus sous le terme "deep learning". Dans la section suivante nous présentons rapidement l'apprentissage profond et comment il peut résoudre notre problème.

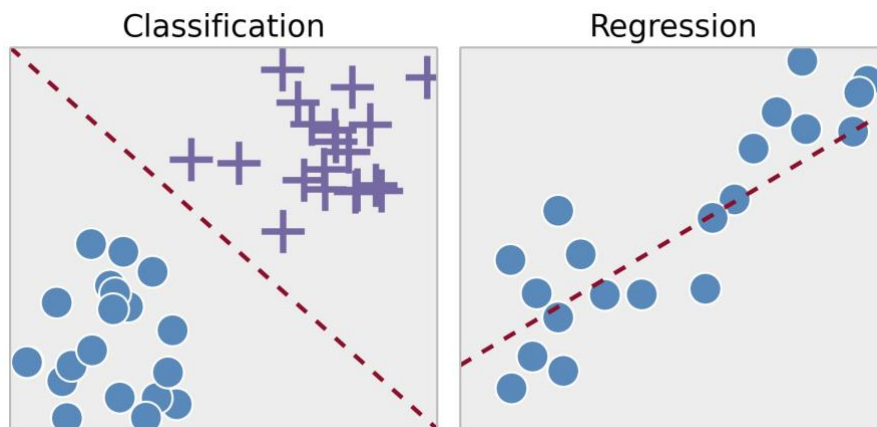


Figure 2 - (Gauche) Représentation d'un problème de classification. (Droite) Représentation d'un problème de régression.

4.2 Le perceptron

Avant l'apprentissage profond, les méthodes d'apprentissage étaient limitées dans leur capacité à traiter des données brutes. Il était nécessaire de trouver une méthode d'extraction de caractéristiques assez efficace permettant au système d'apprentissage de reconnaître les motifs dans les données. L'apprentissage de représentation est un ensemble de technique qui permet à une machine à laquelle on fournit des données brutes de découvrir la représentation nécessaire pour résoudre un problème donné. Enfin les méthodes d'apprentissage profond sont des méthodes d'apprentissage de représentation avec plusieurs niveaux de représentation qui transforment la représentation à un niveau en une représentation à un plus haut niveau de manière un peu moins abstraite. Avec un nombre conséquent de transformations, on est capable d'apprendre des fonctions avec un haut niveau de complexité [6].

L'apprentissage profond est une méthode d'apprentissage supervisé. Pour effectuer cet apprentissage, il est nécessaire de récolter un grand nombre de données, chacune de ces données doit être annotée, on doit donc lui attribuer une valeur. Cette valeur peut être une catégorie lors d'un problème de classification ou un réel lors d'un problème de régression [6].

Le concept de réseau de neurones est tout d'abord venu avec l'envie de modéliser les réseaux de neurones biologiques. Il est depuis devenu un problème d'ingénierie avec comme but d'effectuer des tâches d'apprentissage [7].

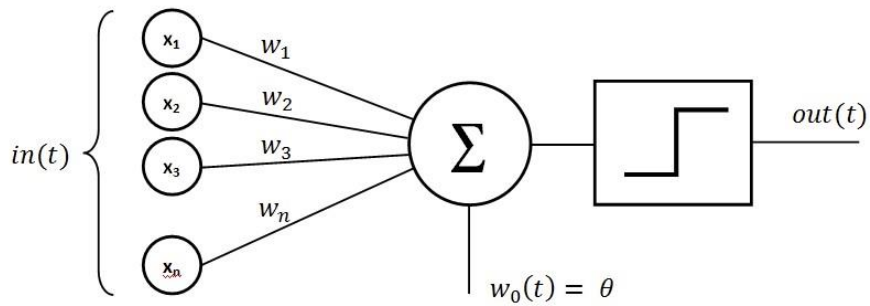


Figure 3 - Le perceptron. Schéma du modèle formel du neurone ou perceptron [8].

Dans la Figure 3, on peut observer le modèle mathématique du neurone qui est le fondement des réseaux de neurones. Les entrées, ici nommées x_n , sont associées à des poids w_n . Ces poids sont appris et contrôlent la puissance d'influence ou d'inhibition d'un neurone sur les suivants [7]. La somme des entrées pondérées par les poids est passée à une fonction non-linéaire appelée fonction d'activation qui détermine la sortie finale du neurone. L'introduction de non linéarité permet au réseau de neurones d'apprendre à faire des prédictions sur des modèles non-linéaires comme il en existe beaucoup dans le monde qui nous entoure.

4.3 Le perceptron multicouches

Ce modèle de neurone a pour intérêt d'avoir la capacité d'être lié à d'autres neurones aussi bien en entrée qu'en sortie afin de former un perceptron multicouche ou MLP, c'est à dire un réseau de neurones. En pratique on va regrouper les neurones en couches. Le modèle le plus commun est le Fully Connected Layer (couche entièrement connectée) où chaque neurone de la couche est connecté à tous les autres neurones des couches adjacentes à lui-même. Par contre les neurones de la même couche ne partagent aucune connexion [7]. On appelle les couches intermédiaires entre l'entrée et la sortie du modèle des couches cachées comme représentées dans la Figure 4 où on peut voir des couches entièrement connectées.

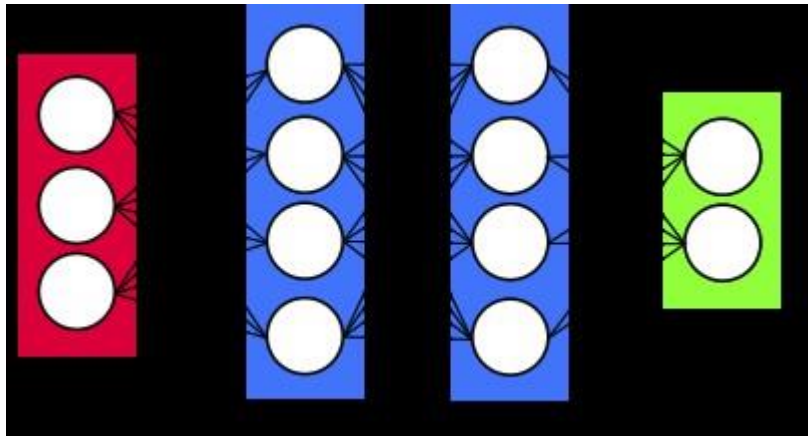


Figure 4 - Architecture d'un réseau de neurone. Un réseau de neurone à trois couches (on ne compte pas la couche d'entrée), deux couches cachées avec chacune 4 neurones et une couche de sortie.

Étant dans un cas d'apprentissage supervisé, le but du réseau de neurones est donc d'être capable de prédire l'étiquette associée à la donnée entrée. En pratique pour la classification, le modèle donne une probabilité de correspondance pour chaque classe. Le but de l'apprentissage est donc de trouver les poids du réseau pour que la classe correspondante à l'entrée obtienne le plus haut score. Pour ce faire, on définit une fonction qui calcule l'erreur entre la prédiction et la vérité. L'apprentissage consiste à minimiser cette fonction appelée fonction de perte ou "loss function" en adaptant les poids du réseau. Dans la pratique le nombre de poids à estimer est très grand et implique qu'un très grand nombre d'exemples est nécessaire pour l'apprentissage [6].

L'apprentissage profond est un champ de recherche très actif. Ses principes de base sont utilisés et adaptés pour adresser une large étendue de problèmes. Il existe ainsi d'innombrables architectures plus ou moins efficaces destinées à l'analyse d'images, de texte, de voix ou encore de vidéos. Dans les parties suivantes nous étudions certains de ces modèles dans le but de les utiliser pour développer notre propre modèle.

5 État de l'art des méthodes d'apprentissage profond

5.1 Les réseaux de neurones convolutifs

5.1.1 Les réseaux de neurones spatiaux convolutifs

Si les réseaux de neurones profonds étaient déjà capables de traiter des images en les représentant sous forme de tableau à une dimension, certains chercheurs se sont penché sur le problème du développement d'un nouveau modèle capable d'analyser des données représentable via des tableaux à multiples dimensions et donc plus adapté pour des images et particulièrement lorsque les images sont de grande taille. À l'origine, la première mention de convolution pour les réseaux de neurones apparaît sous le nom de Neocognitron [9]. Mais les CNNs (Convolutional Neural Networks) ont réellement été formulés dans [10, 11]. Ils sont en partie responsables de la popularisation de l'apprentissage profond en obtenant de bons résultats là où les réseaux de neurones classiques peinaient. Ils sont maintenant complètement adoptés dans le champs de la vision par ordinateur [6].

Les réseaux de neurones convolutifs introduisent de nouvelles couches dans les réseaux de neurones en apportant notamment des couches à 3 dimensions. Les blocs de constructions élémentaires pour l'architecture d'un réseau de neurones convolutifs sont :

- Les couches de convolution (CONV) qui traitent les données localement d'un champ récepteur.
- Les couches de pooling (POOL), qui permettent de compresser les informations et d'amener de l'invariance en translation.
- Les fonctions d'activation
- Les couches entièrement connectées (FC) similaires à celles des MLPs souvent présentes en sortie du réseau.

Les couches de convolution sont la base d'un CNN. Le but est de focaliser les différents neurones sur des zones locales de l'image, à l'instar des filtres lors de convolution. Les poids des neurones jouent le rôle des coefficients du filtre. Pour des images de dimension 3 (hauteur x largeur x 3), ces coefficients sont mis sous la forme d'une matrice 3D. Une couche de convolution est définie en précisant la taille des filtres (identique pour toute la couche) et le nombre de filtres, comme on peut le voir sur la figure 5. Des éléments de réglage sont également nécessaires comme le stride qui détermine le décalage du filtre entre chaque zone de l'image. Un stride de 1 décale le filtre de un pixel tandis qu'un stride de 2 décale le filtre de 2 pixels et amène donc à une sortie deux fois plus petite que l'entrée. Enfin le padding permet d'ajouter (ou non) des zéros sur les bords de l'image afin de contrôler la taille des sorties de la convolution. En chaque pixel de l'image, autant de caractéristiques sont extraites que le nombre de filtres [12].

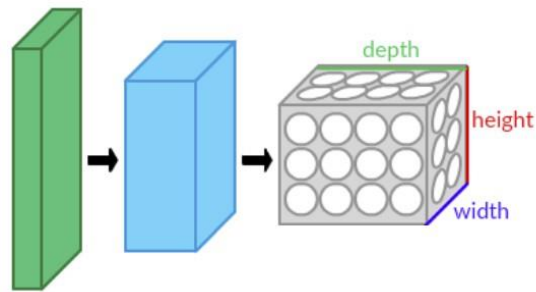


Figure 5 - Couche du CNN en 3 dimensions. Le vert est le volume d'entrée, le bleu le volume du champ récepteur et le gris la couche de CNN. Les cercles représentent les neurones artificiels indépendants [13].

On ne va pas rentrer en détail dans le fonctionnement des réseaux de neurones convolutifs mais ce qu'il faut retenir c'est qu'ils sont particulièrement efficaces pour apprendre les caractéristiques des images. On peut observer leur fonctionnement dans la Figure 6.

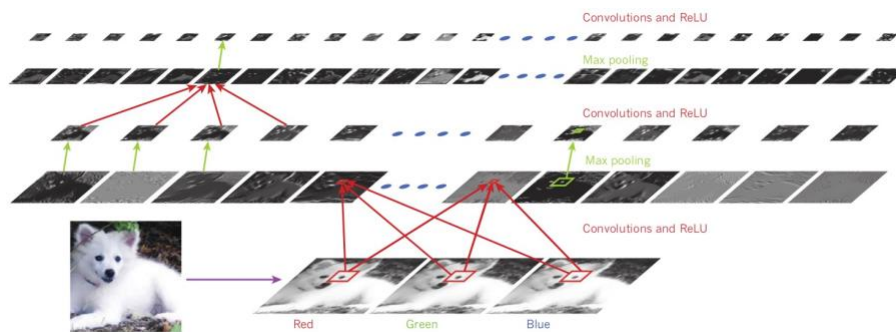


Figure 6 - À l'intérieur d'un réseau de neurones convolutifs. Chaque image est une carte des caractéristiques correspondant à la sortie d'une couche. L'information est de plus en plus complexe à mesure qu'on accumule les couches, les couches basses apprenant les caractéristiques basiques comme les contours [6].

5.1.2 Les réseaux de neurones temporels convolutifs 3D

Avant le modèle C3D dont nous allons parler dans la partie suivante, nous introduisons le concept de réseaux de neurones temporels convolutifs 3D abordés dans certains articles comme [14, 15, 16]. Les CNNs, tels qu'introduits dans la partie précédente, utilisent la convolution pour analyser des images avec 1 ou plusieurs canaux couleur (Figure 7). Dans le cas de vidéos, la dimension temporelle doit être gérée de manière spécifique de manière à être conservée (ce qui n'est pas le cas du nombre de canaux pour une image couleur). L'idée des réseaux de neurones temporels convolutifs 3D est de parvenir à éviter la perte de cette information afin d'analyser des séquences temporelles.

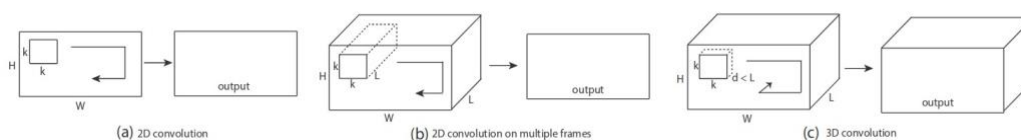


Figure 7 - Opérations de convolution 2D et 3D. a) Appliquer une convolution 2D sur une image donne une image. b) Appliquer une convolution 2D sur une vidéo représentée comme un volume donne aussi une image. c) Appliquer une convolution 3D sur une vidéo représentée comme un volume donne un volume préservant les données temporelle du signal d'entrée [16].

5.1.3 Le modèle C3D

En 2014, D. Tran et al. proposent une nouvelle solution pour l'apprentissage des caractéristiques spatio-temporelles en utilisant des réseaux temporels convolutifs 3D [16]. La proposition a pour but d'être une solution générique, compacte, efficace en calculs et simple à implémenter. Cette méthode reprend le fonctionnement des réseaux convolutifs 2D. Le principal intérêt est de conserver les données temporelles après la convolution contrairement aux réseaux convolutifs 2D comme on peut le voir dans la figure 7. Le modèle C3D a été appris sur de courtes séquences de vidéo (16 images) et peut-être appliqué à la reconnaissance d'action humaine (HAR : Human Action Recognition).

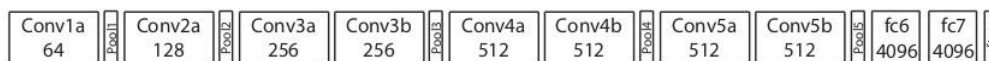


Figure 8 - Architecture C3D. Composé de 8 convolutions, 5 max-pooling and 2 couches entièrement connectées suivies d'une couche de sortie softmax. Tous les noyaux convolutifs 3D sont de taille 3 x 3 x 3 avec un stride = 1 dans les dimensions spatiale et temporelle [16].

D'après [16], le modèle C3D se positionne parmi les méthodes de pointes pour la reconnaissance d'action mais se différencie surtout en étant efficace, compacte et simple d'utilisation.

5.2 Les réseaux de neurones récurrents

5.2.1 Les RNNs

Un des défauts majeurs des réseaux de neurones à propagation avant ou FFNNs (Feed Forward Neural Networks) est leur incapacité à traiter des données dépendantes les unes des autres. En effet à chaque fois qu'une donnée est traitée, le réseau ne conserve pas les informations décrivant cette donnée. Cette contrainte empêche le traitement de données telles que les vidéos, l'audio ou encore les séquences de mots. De plus, les réseaux de neurones standards utilisent souvent en entrée des vecteurs de taille fixe rendant difficile l'analyse de séquences à durée variable [17].

Les réseaux de neurones récurrents ou RNNs sont une classe de modèle qui a la capacité de transmettre des informations à travers les séquences tout en traitant les données

séquentielles les unes après les autres comme un réseau de neurone standard. Ils se basent sur les FFNNs mais incluent des liens entre les neurones d'une même couche qui se transmettent les données des pas temporels successifs comme sur la Fig. 9.

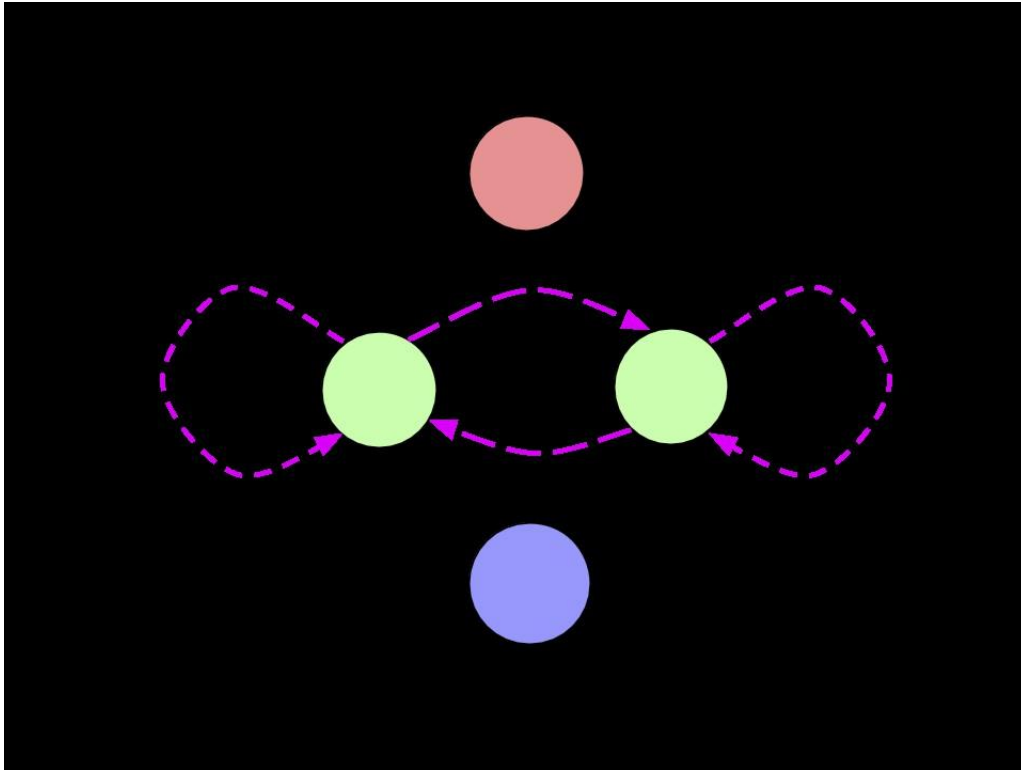


Figure 9 - Un RNN simple. À chaque pas temporel l'activation est transmise selon les flèches continues comme dans un FFNN. Les flèches discontinues connectent un noeud source à chaque temps t à un noeud cible à un temps $t + 1$ [17].

On peut représenter la dynamique du réseau au cours des différents pas temporels dans la Fig. 10.

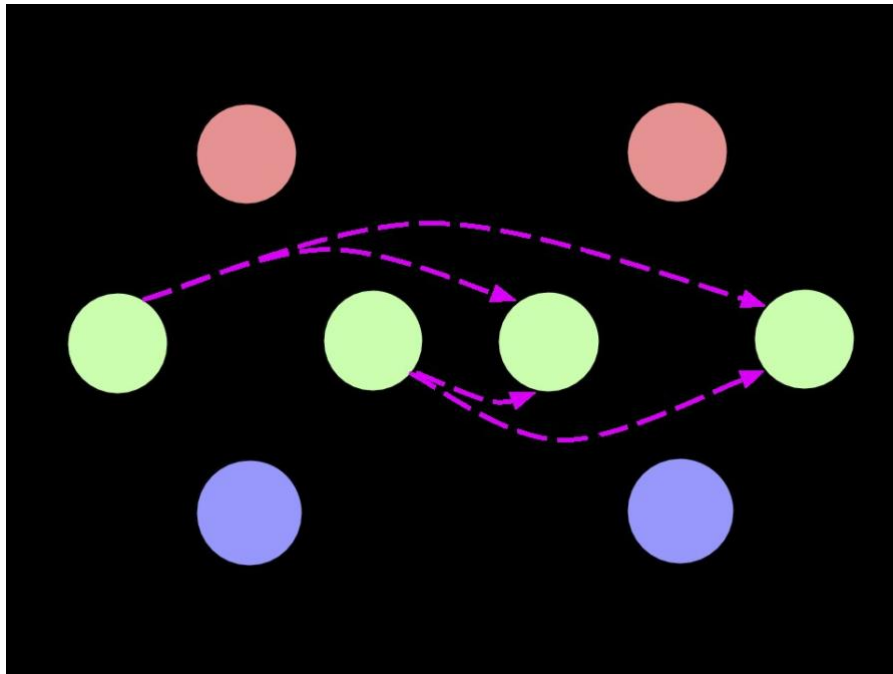


Figure 10 - Le RNN de la Fig. 9 déployé au cours des pas temporels. [17].

5.2.2 Les LSTMs

Afin de faire face au problème de la disparition des gradients sont introduits en 1997 les réseaux de neurones à mémoire long et court terme ou LSTMs (Long Short-Term Memory) [18]. L'idée est d'instaurer un modèle qui soit capable d'apprendre les caractéristiques temporelles pertinentes qu'elles soient espacées aussi bien par de courtes que de longues périodes.

Dans [18] est décrit l'architecture LSTM qui est composée d'une suite d'unités LSTM. À l'intérieur de ces unités on va retrouver des portes qui ont pour but d'indiquer quelles informations doivent être gardées et lesquelles doivent être oubliées. C'est fait en modifiant l'état de la cellule comme on peut l'observer sur la figure 11.

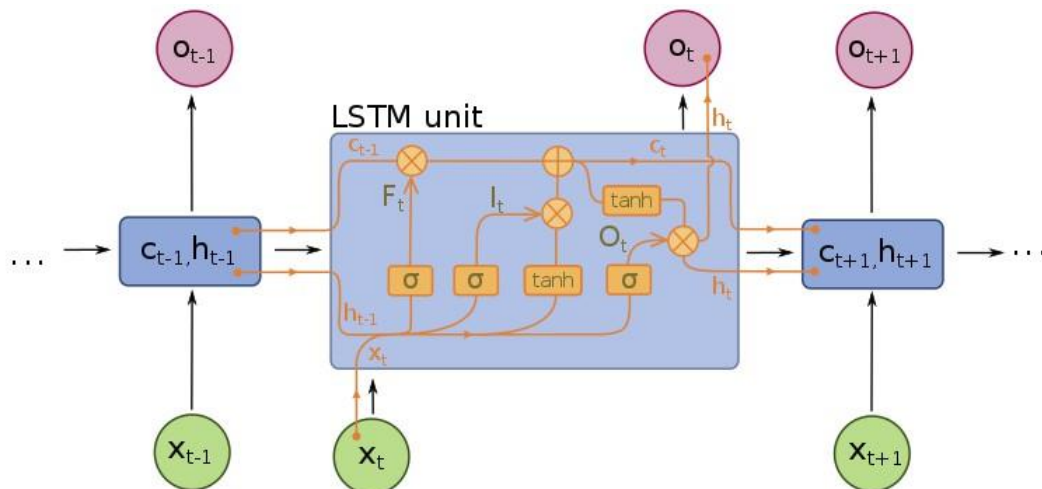


Figure 11 - Une cellule LSTM. [19]

Le modèle plus complexe inclut notamment une porte d'oubli (forget gate) qui permet, via un ensemble de poids, d'indiquer quelles sont les informations qui doivent être conservées et celles qui doivent être oubliées. Ces portes sont composées d'une fonction d'activation, ici tanh ou sigmoïde, et leurs poids sont déterminés lors de la phase d'apprentissage.

Le modèle LSTM est un standard lorsqu'il est nécessaire d'analyser des données temporelles et qu'un RNN simple ne suffit pas. Par la suite certains chercheurs ont apporté des variantes à ce modèle notamment en changeant l'architecture de la figure 11 mais ces méthodes ne sont pas détaillées ni utilisées dans ce projet.

6 Exposé des travaux

Dans cette partie nous présentons le travail accompli lors de la période de stage. Le but est de créer un modèle capable d'évaluer le mouvement général spontané d'un enfant. Différents types de mouvements peuvent amener une séquence à être classée comme anormale, néanmoins l'ensemble de données fourni contient en grande majorité des séquences de type pauvre répertoire. La catégorie pauvre répertoire est caractérisée par 5 facteurs :

- Des séquences de mouvement répétitives ;
- Une faible variabilité ;
- Une faible complexité ;
- Une vitesse monotone ;
- Une fluidité assez faible.

Pour estimer ces paramètres il est d'abord nécessaire d'avoir une bonne estimation de la pose de l'enfant afin d'évaluer la complexité qui analyse les caractéristiques du mouvement en lui même. Mais il faut aussi une bonne évaluation de leur évolution dans le temps pour repérer la vitesse monotone et la faible fluidité. Pour analyser la variabilité et la répétitivité il faut une bonne estimation de la pose et du mouvement ainsi que leur évolution au cours du temps. L'idée d'apprentissage automatique des vidéos est de caractériser le mouvement général spontané directement, sans contraindre la pose des enfants.

Cette section se découpe en trois parties, chacune décrivant un modèle. La première méthode fait appel au modèle C3D couplé à un réseau LSTM, la deuxième méthode utilise des convolutions temporelles 1D et enfin la troisième méthode analyse le flot optique des vidéos de plusieurs façons. Chaque section s'articule en décrivant d'abord le modèle, ensuite les réalisations effectuées pour poursuivre avec les expérimentations et les résultats obtenus avant de conclure sur ce modèle.

6.1 1^{er} Modèle : Analyse des caractéristiques spatio-temporelles via une approche double utilisant un modèle C3D et un LSTM

Pour résoudre le problème énoncé il est donc nécessaire d'être en capacité d'évaluer la pose et le mouvement de l'enfant de manière précise mais aussi de caractériser son évolution au cours du temps.

Précédemment nous avons introduit le modèle C3D, un réseau de neurones temporel convolutif 3D capable de reconnaître des actions sur des périodes de 16 images. Ce modèle atteint des performances de pointe dans le domaine de la reconnaissance d'actions humaines [16]. On émet donc l'hypothèse que l'utilisation de ce modèle pourrait être efficace concernant la caractérisation des mouvements généraux spontanés. Néanmoins, en l'état, s'appliquant seulement sur une courte période, le modèle est incapable d'évaluer la répétitivité et la variabilité des séquences de GMs.

Par ailleurs, nous avons aussi introduit les LSTMs, des réseaux de neurones récurrents ayant la capacité de prendre en compte efficacement des informations à court et long terme pour la caractérisation de séquences.

On retrouve dans la littérature scientifique de multiples exemples qui combinent ces deux approches pour une analyse précise du mouvement à court et long terme [20, 21, 22]. Dans cette partie nous proposons une approche en deux étapes : premièrement le modèle C3D caractérisera le mouvement de morceaux des vidéos en créant des descripteurs, ces descripteurs seront ensuite fournis à un modèle LSTM qui sera entraîné à estimer l'évolution du mouvement au cours du temps.

Dans la partie suivante nous détaillons le modèle mis en place, ensuite nous expliquons l'implémentation du modèle ainsi que les expérimentations induites. Enfin nous présentons les résultats obtenus et concluons sur l'utilisation de ce modèle.

6.1.1 Modèle

Dans un premier temps il est nécessaire de déterminer la manière dont nous allons utiliser C3D. En effet C3D décrit seulement un modèle, c'est à dire un ensemble de couches rappelés dans la Fig. 8.

Lorsqu'on est confronté à l'utilisation d'un modèle déjà existant plusieurs possibilités s'offrent à nous. Premièrement on peut choisir d'entraîner complètement le modèle avec les données à notre disposition, cela nécessite d'avoir un grand nombre de données et peut prendre beaucoup de temps surtout si on cherche à obtenir des performances optimales. L'avantage est que le modèle est complètement adapté au problème traité. On peut aussi choisir d'utiliser une version de ce modèle qui a déjà été entraînée sur un ensemble de données différentes. On récupère alors l'ensemble des poids déterminés à l'issue de l'entraînement. Si le modèle est populaire il est fort probable que cette ressource soit mise en ligne. On peut donc choisir d'utiliser ce modèle sans avoir besoin de l'entraîner sur nos données. Il existe aussi une approche intermédiaire qu'on appelle le fine-tuning qui consiste à charger les poids entraînés sur un ensemble de données différent et à poursuivre l'entraînement avec nos données personnelles. Quand l'ensemble de données est assez grand et quand les caractéristiques de nos données sont proches des données utilisées pour le premier entraînement le fine-tuning peut mener à de bons résultats.

Le choix dépend alors du contexte applicatif. Dans notre cas on possède environs 800 vidéos d'entraînement. Étant donné la complexité de l'information à extraire et le coût en temps d'apprentissage, l'apprentissage complet du modèle n'a pas été retenu pour notre application.

Pour les deux autres possibilités il est nécessaire de trouver un modèle entraîné sur ensemble de données existantes. La base de données SPORTS-1M [15] répertorie 1 million de vidéos de sport issues de YouTube appartenant à 487 catégories. Dans [16] D. Tran et al. entraînent leur modèle sur cet ensemble de données. Étant donné sa capacité à caractériser les mouvements sportifs, on peut émettre l'hypothèse qu'il est aussi capable de reconnaître les mouvements des enfants. Il est donc nécessaire de choisir entre une utilisation en l'état ou un fine-tuning. Étant donné que les annotations que nous avons décrivent les séquences, effectuer un fine-tuning (ou même un entraînement à partir de rien) implique de créer un modèle hybride regroupant C3D et le LSTMs en un seul gros modèle. Cela implique de très longs temps de calculs (plusieurs semaines d'apprentissage). Nous n'avons pas retenu cette option pour ce stage court de 2 mois. Il reste ainsi

l'utilisation du modèle en l'état qui sera utilisé pour l'extraction de caractéristiques (ou features extraction).

L'extraction de caractéristiques consiste à ne pas utiliser la totalité du modèle. Ainsi, nous utilisons seulement les premières couches du réseau qui permettent de représenter la séquence d'entrée par un vecteur de caractéristiques. L'hypothèse émise ici est que ce descripteur de 16 images consécutives permet de décrire le mouvement du bébé et est assez fin pour que leur séquence le long de la vidéo caractérise les GMs.

Donc, chacune de nos vidéos passe à travers le modèle C3D entraîné sur la base de données SPORTS-1M. Les vidéos sont donc découpées en séquences de 16 images sur lesquelles une caractérisation est extraite avec le modèle C3D. On obtient donc une suite temporelle de descripteurs qui est mise en entrée d'un modèle LSTM. Celui-ci analyse l'évolution temporelle des descripteurs au cours du temps et cherche à prédire l'anormalité des séquences. La sortie du réseau sera binaire : séquence vidéo normale ou anormale.

Dans cette partie nous avons donné une description globale du modèle et explicité les choix fondamentaux dans sa conception. Dans la prochaine partie nous détaillons l'implémentation de ce modèle.

6.1.2 Implémentation du modèle

Pré-traitement. Il est commun dans les tâches d'apprentissage de ne pas utiliser les données en l'état et d'appliquer des actions de pré-traitement afin d'assurer que les données soient comparables entre elles mais aussi qu'elles soient adaptées au modèle utilisé, le modèle C3D dans notre cas.

Les vidéos de notre base de données ont une fréquence d'acquisition bien plus importante que celle des vidéos du modèle C3D. Comme le modèle C3D analyse les vidéos par séquence de 16 images il est important que ces 16 images correspondent à la même durée pour toutes les vidéos. Pour obtenir un maximum d'informations pendant la période de 16 images, tout en ne réduisant pas trop la fréquence d'image afin de garder du détail, nous avons ré-échantillonné les vidéos à une fréquence de 25 images par seconde.

Ensuite il est nécessaire de redimensionner et de recadrer les vidéos car C3D prend comme entrée des vidéos de taille 16x171x128x3 qui sont rognées en 16x112x112x3 (16 représente le nombre d'image, 112x112 la taille de chaque image et 3, le nombre de canaux couleur). Comme dans l'implémentation fournie par [16], on soustrait à chaque bloc de 16 images la moyenne calculée sur l'ensemble d'entraînement SPORTS-1M.

Le modèle C3D impose de fournir des séquences de 16 images, il est donc nécessaire de découper nos vidéos. À l'instar de [16] nous effectuons un chevauchement de 8 images entre les séquences de 16 images. En effet cette redondance d'information peut mettre en avant certains mouvements ou changements de rythme et permet également de limiter la discontinuité entre les segments.

Extraction des descripteurs C3D. Une fois les images prêtes à être traitées, le modèle C3D est utilisé pour extraire les caractéristiques représentées sous la forme d'un vecteur de dimension 4096. Pour chaque vidéo, ces descripteurs sont regroupés au cours du temps dans une matrice.

Pour cela il suffit de retirer la couche de softmax ainsi que le dropout qui de toute façon n'a pas d'importance hors de la phase d'entraînement. Un fois ce descripteur de taille 4096 récupéré on va le ranger dans un tableau avec tous les autres descripteurs de la vidéo. Ainsi on passe d'une séquence d'image (la vidéo d'origine) à une séquence de descripteur du mouvement.

Pré-traitement avant le LSTM. Une fois l'ensemble des données obtenues, elles sont standardisées c'est à dire qu'on calcule la moyenne et l'écart-type de l'ensemble des données et ensuite on soustrait à chaque donnée la moyenne et la divise par l'écart-type, ce qui permet de traiter les séquences de manière équivalente. De plus il est nécessaire de mettre en forme l'ensemble de données afin qu'à chaque vidéo corresponde une étiquette. Dans le cas de la classification binaire la vidéo est classée comme normale ou anormale et dans le cas de la régression on reporte le score déterminé par les médecins. Une fois cet ensemble déterminé on peut passer à l'entraînement. L'ensemble de données est divisé en 3 parties :

- l'ensemble d'entraînement qui sera dédié à l'entraînement du réseau.
- l'ensemble de validation qui permet de vérifier la prédiction sur un autre ensemble pendant l'entraînement et ainsi constater qu'il n'y a pas de sur-apprentissage. Cet ensemble sert aussi dans certaines méthodes de régularisation afin d'éviter le sur-apprentissage
- l'ensemble de test qui permet d'avoir une réelle estimation des performances de notre modèle sur des données n'ayant eu aucun contact avec l'entraînement et donc de valider si le modèle fonctionne correctement.

Choix du modèle LSTM. Il est difficile de déterminer a priori un modèle approprié à un problème sans effectuer de tests. Tout le processus d'entraînement s'effectue en modifiant certains paramètres afin d'obtenir le meilleur résultat possible. On commence avec un modèle basique que l'on adapte ensuite. Nous choisissons d'effectuer d'abord la classification, le but ici est de déterminer si une séquence est normale ou anormale.

Entraînements

Suivi d'entraînement Lors de l'entraînement d'un réseau de neurone il est important de suivre son évolution pour comprendre ce qui peut faire défaut. Il existe plusieurs données que l'on peut observer. Dans ce rapport, on se contente d'observer la fonction perte et le taux de reconnaissance sur l'ensemble d'entraînement et sur l'ensemble de validation.

Entraînement naïf On commence donc par un modèle simple afin d'observer les premiers résultats. Ce modèle contient des couches LSTM et des couches FC comme présenté Fig. 12. Il est entraîné sur 20 epochs, c'est à dire qu'on fait passer 20 fois l'ensemble des données d'entraînement à travers le réseau de neurones. C'est suffisant pour voir s'il peut être intéressant d'entraîner le modèle sur plus d'epochs ou non. Pour

commencer on découpe les données d'entraînement en batch de 32 éléments, cette valeur pourra être changée plus tard.

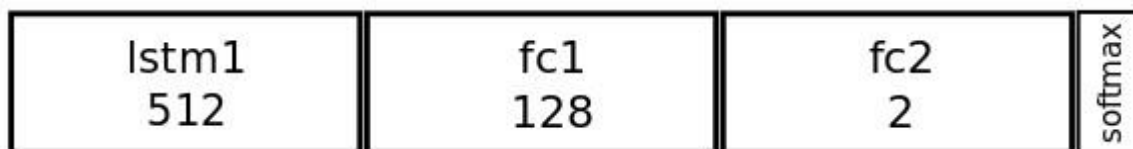


Figure 12 - Modèle LSTM naïf avec une couche de LSTM avec un vecteur latent de dimension 512 et une couche entièrement connectée avec 128 neurones.

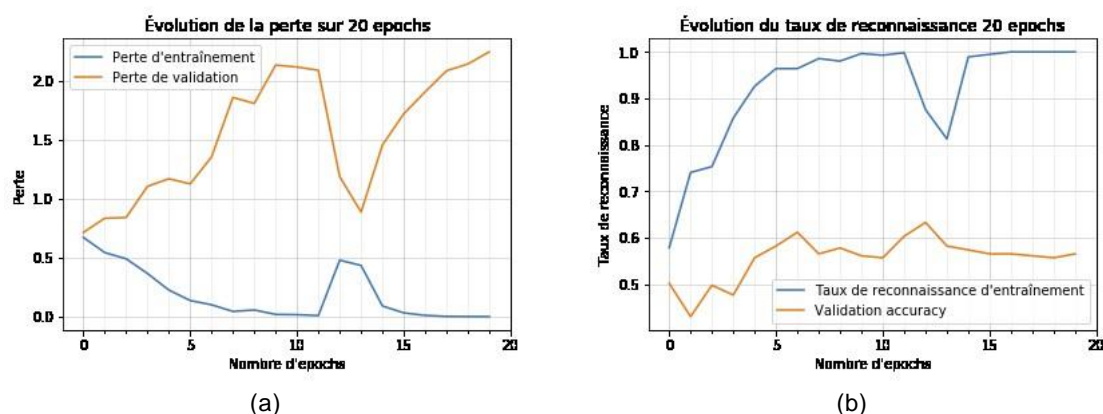


Figure 13 - Entraînement du modèle LSTM naïf . (a) Graphique représentant l'évolution de la fonction perte (b) Graphique représentant l'évolution du taux de reconnaissance.

	Normal prédit	Anormal prédit
Normal véritable	62	49
Anormal véritable	54	72

Figure 14 - Matrice de confusion du modèle LSTM naïf .

Dans la figure 13, on observe que l'ensemble d'entraînement obtient rapidement un taux de reconnaissance de 1 et une diminution de la fonction perte, a contrario de l'ensemble de validation. On est donc dans un cas de sur-apprentissage, le modèle a mémorisé les entrées et fait donc une prédiction parfaite sur l'ensemble d'entraînement, par contre dès qu'il est confronté à de nouvelles données il effectue sa prédiction au hasard comme le montre la matrice de confusion représentée en figure 14. En plus de faire du sur-apprentissage, on remarque que même si le taux de reconnaissance augmente légèrement pour l'ensemble de validation, la fonction perte ne fait qu'augmenter. On peut donc en déduire que le modèle n'est pas efficace. Comme il est capable de faire du sur-apprentissage c'est que le

réseau est capable de différencier les séquences mais qu'il est incapable de généraliser. On peut penser donc qu'il n'est pas capable de voir les caractéristiques qui sont inhérentes aux GMs anormaux. Pour ça on peut donc modifier les options de régularisation mais aussi les différents hyperparamètres, c'est à dire les paramètres non appris par le réseau de neurones mais fixé a priori.

Réduction du sur-apprentissage Pour ce nouveau modèle on présente celui qui a obtenu les meilleurs résultats parmi nos nombreux tests. Il existe plusieurs façons pour réduire le sur-apprentissage. Le dropout est une méthode qui consiste à empêcher le réseau de neurone de prendre certaines branches qui relie les différentes couches. On donne la proportion d'unité qu'on veut empêcher d'être entraînée à chaque mise à jour du réseau de neurone. On peut aussi utiliser la régularisation sur les poids qui permet de limiter leur valeur et donc de les garder relativement petits. Des poids trop grands mènent à un réseau instable avec une grande probabilité de sur-apprentissage [23].

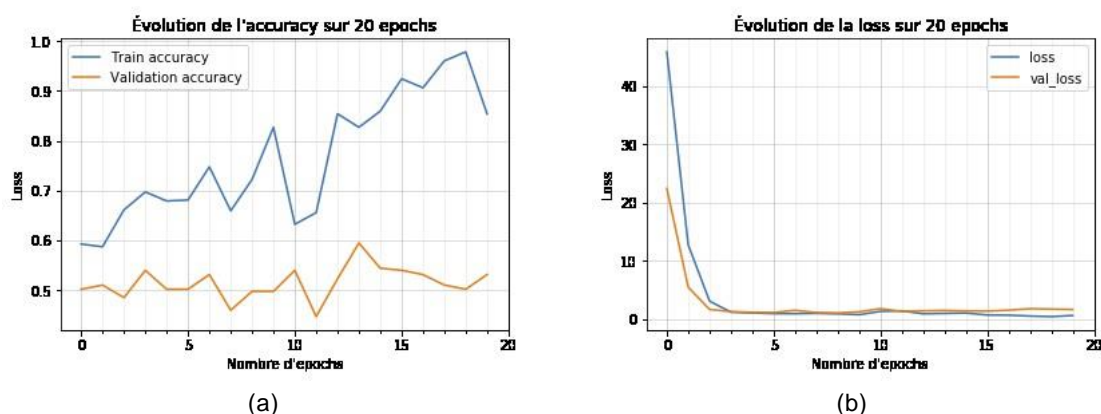


Figure 15 - Entraînement sur un modèle LSTM avec réduction du sur-apprentissage. (a) Graphique représentant l'évolution du taux de reconnaissance du modèle (b) Graphique représentant la perte du modèle.

	Normal prédit	Anormal prédit
Normal véritable	0	111
Anormal véritable	0	126

Figure 16 - Matrice de confusion du modèle LSTM avec réduction du sur-apprentissage.

On observe avec les figures 15 et 16 qu'avec les options de régularisation on obtient un réseau qui n'apprend pas du tout. Dans la figure 16 on voit que le modèle classe les vidéos comme anormales dans tous les cas. De plus dans la figure 15 on observe que la fonction perte converge en ≈ 1.6 et ne montre pas d'indice indiquant qu'elle va redescendre. Lors d'autres entraînements on arrive à réduire encore plus le sur-apprentissage mais les résultats du taux de reconnaissance sont toujours équivalents au hasard.

Avec ce modèle utilisé en l'état nous n'avons pas réussi à trouver des résultats capables de déterminer avec un pourcentage satisfaisant la qualité des GMs. On peut émettre des hypothèses sur la raison du non fonctionnement de ce modèle. L'ensemble de données est construit de telle sorte que pour chaque phase (A : pré-terme, B : terme, C : 15 semaines après terme), chaque enfant possède 1 vidéo ou n'en possède pas. Cette vidéo obtient un score et une classe (normale ou anormale) elle est ensuite découpée en un certain nombre de séquences pour visualiser uniquement les moments où on observe des GMs. Ainsi chaque séquence obtient le même score ou la même classe ce qui implique des caractéristiques communes au-delà des mouvements (l'enfant et l'environnement). On peut donc imaginer que cette redondance mène le réseau à mémoriser rapidement car les caractéristiques autres que les GMs sont plus évidentes lorsqu'on regarde les séquences de descripteurs. D'autre part, la dimension de l'entrée est très importante et peut également conduire au sur-apprentissage. Dans la partie suivante, on va considérer ce problème.

6.1.3 Réduction de dimension et traitements additionnels

Afin de diminuer la taille des caractéristiques on introduit l'analyse en composante principale ou ACP. L'ACP est une technique de réduction de dimension. La réduction de dimension consiste à transformer un lot de données qui est exprimé selon un certain nombre de dimension (ici 4096) et les réduire tout en conservant un maximum d'information. Cela peut permettre de réduire un lot de données en 2 ou 3 dimensions pour le visualiser dans un graphique ou d'améliorer les performances d'un modèle en réduisant la taille des informations.

L'ACP est une technique de réduction de dimension qui privilégie les données de variance maximale. Pour résumer l'ACP consiste en des transformations géométriques qui projettent les données sur l'axe des variances maximales. Comme on souhaite conserver les données qui varie, c'est à dire les mouvements et les changements dans la position de l'enfant, et qu'on souhaite ignorer les aspects redondants, éléments codant l'environnement comme le berceau par exemple, l'ACP peut être une solution adaptée. Le côté négatif de cette technique est que l'on ne connaît pas vraiment les informations que l'on traite (descripteurs C3D) et donc, conserver les informations de variance maximale peut aussi bien enlever des informations intéressantes.

De plus afin de faciliter la reconnaissance, une rotation a été appliquée à toutes les vidéos afin que l'enfant ait la tête vers la droite mais aussi un recadrage afin d'éviter de voir les mouvements du personnel autour du berceau. On gardera les vidéos sous cette forme à partir de maintenant.

Malgré l'intérêt que porte cette méthode pour d'autres tâches d'apprentissage, l'ACP n'a donné pas amélioré les résultats pour notre application.

Dans un second temps afin de simplifier le problème, nous avons utilisé des données d'une seule phase afin d'avoir des données plus comparable. En effet, les gestes d'un enfant pré-terme et d'un enfant à 15 semaines après le terme sont différents. On choisit la phase B pour nos tests car les enfants sont plus actifs et ont des mouvements plus expressifs que la phase A et qu'on a un plus grand nombre de séquences que dans la phase C. Dans les modèles suivants on utilisera uniquement cet ensemble de données.

Là aussi, malgré cette simplification, les résultats n'ont pas été concluants.

6.1.4 Conclusion

Dans cette section nous avons mis en place un modèle double alliant les forces du modèle C3D pour ce qui est de l'extraction des attributs et les forces du modèle LSTM pour évaluer des séquences à court et long terme. Nous avons modifié le modèle afin qu'il obtienne les meilleurs scores. Nous avons utilisé une méthode de réduction de dimension afin de limiter l'influence des caractéristiques de faible variance et avons utilisé un ensemble de données réduits transformés afin d'aider le modèle à généraliser. Enfin nous avons réduit l'ensemble de données pour que les enfants aient tous un âge appartenant à la phase B. Tous ces éléments n'ont pas permis d'obtenir des résultats probants concernant la prédiction des GMs. On peut émettre plusieurs hypothèses sur les raisons de cet échec.

L'ensemble de données d'apprentissage est peut-être trop petit. En effet, les mouvements sont très variés et reconnaître quelques spécificités liées à l'anormalité est très complexe. Ainsi, la base ne présente peut-être pas une assez grande diversité de mouvement pour que l'apprentissage puisse repérer ceux significatifs d'une anormalité.

On peut aussi se demander si le modèle C3D entraîné sur des vidéos de sports est adapté pour assimiler les détails des mouvements des enfants. En parcourant les documentations officielles de demande d'étude observationnelle, on comprend que les extrémités comme les pieds ou les mains sont des éléments importants. Les séquences de sports étant plus caricaturales la caractérisation de C3D n'est peut-être pas assez fine pour notre application.

À la suite de ces mauvais résultats et dû à des contraintes de temps, nous avons choisis d'essayer d'implémenter d'autres modèles. Il faut noter que ces résultats ne sont cependant pas suffisants pour conclure sur l'inefficacité du modèle C3D. Pour aller plus loin on pourrait par exemple choisir une approche plus hybride, c'est à dire créer un modèle fusionnant C3D et LSTM dans le même modèle et faire de l'ajustement ou un entraînement à partir de rien de tout le modèle, même si notre ensemble de données est sans doute trop petit pour avoir des résultats correct eu égard au nombre de paramètres à estimer.

Dans la partie suivante nous proposons une nouvelle approche qui encode une séquence en un vecteur décrivant les actions de l'entrée. Cette approche va être appliquée aux séquences de descripteurs C3D.

6.2 Encodeur TCN

Dans [24] Lea et al. introduisent une nouvelle classe de modèle, les réseaux convolutifs temporels ou TCNs (Temporal Convolutional Networks) qui utilisent des convolution temporelles afin d'effectuer de la segmentation ou de la détection d'actions. Ils décrivent un modèle Encoder-Decoder TCN et un modèle Dilated TCN qui capturent de manière différente les motifs sur une longue portée. Ils montrent que ces TCN sont capables de capturer les compositions d'actions, la durée des segments, et les dépendances sur le long terme et sont bien plus rapides à entraîner que les LSTMs [24].

6.2.1 Modèle

Dans notre cas on s'intéresse au modèle Encoder-Decoder TCN représenté dans la figure 17. En l'état, le modèle fait une prédiction sur chaque instant temporel afin de faire de la segmentation et de la détection. Notre problème est différent, on souhaite décrire chaque séquence pour la classifier comme normale ou anormale ou pour lui attribuer un score de normalité. Afin d'exploiter les performances de ce modèle et de les adapter à notre problème on choisit de conserver uniquement la partie encodeur. En effet les valeurs issues de l'encodage sont par définition des éléments codant la description temporelle de la séquence. Pour ça on combine la sortie de l'encodeur avec un MLP. En somme les vidéos sont encodées une première fois en séquences de descripteurs par C3D puis encodées en un vecteur latent par le TCN. Le MLP final est utilisé pour la classification.

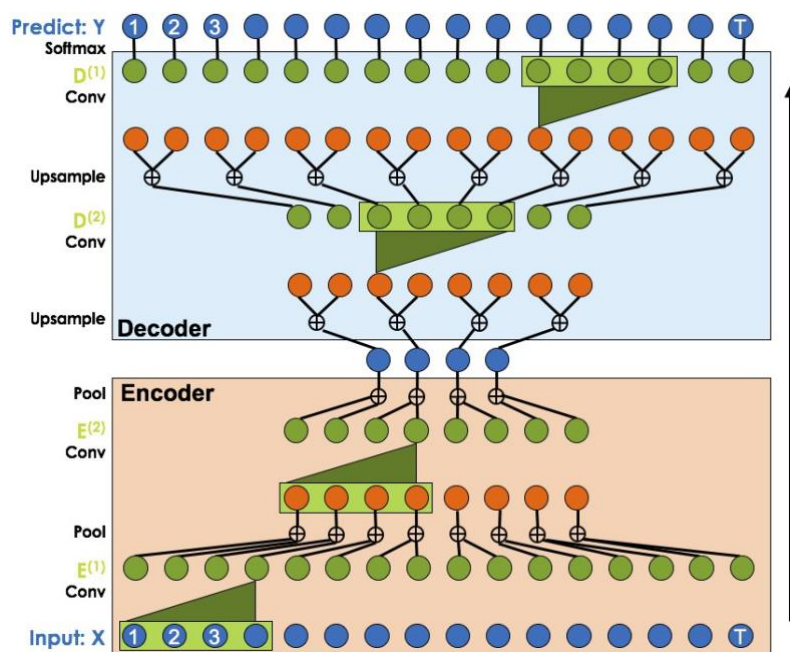


Figure 17 - Encoder-Decoder Temporal Convolutional Network (ED-TCN). Le modèle utilise des convolutions temporelles, du regroupement (pooling) et de l'augmentation (upsampling) [24].

6.2.2 Implémentation

On ajoute donc au modèle 2 couches FC et 1 classifieur. On entraîne le modèle sur l'ensemble des descripteurs issus des vidéos de la phase B qui ont été recadrées et ont subi une rotation. En observant les résultats dans la Fig. 18 on observe un overfitting rapide sur le modèle.

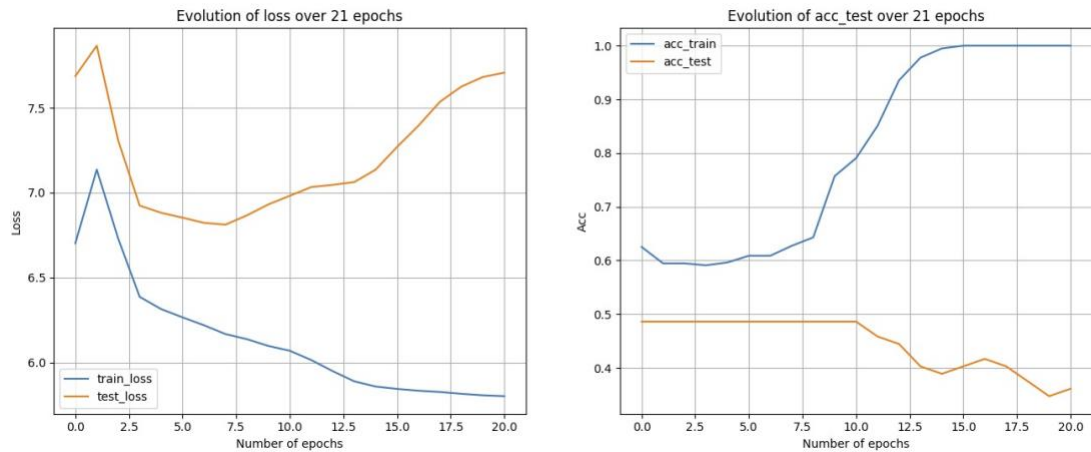


Figure 18 - Loss et accuracy du modèle encodeur TCN.

6.2.3 Conclusion

Dans cette partie nous implémenté une méthode différente des LSTMs pour la description de séquences temporelles en combinant un aspect des TCN et un MLP classique. Les premières expérimentations n'ont pas été concluantes mais l'idée d'utiliser ce modèle pour les descriptions de séquences temporelles paraît intéressante. Encore une fois, ce modèle repose sur l'hypothèse que le modèle C3D décrit bien les séquences de GMS, ce qui n'est pas garanti. Ainsi, si les séquences sont mal décrites, le TCN sera incapable de capter les informations pertinentes.

Il serait donc intéressant de tester cette méthode avec d'autres moyens pour encoder les séquences de mouvements que C3D.

Dans la partie suivante, nous avons décidé de faire des prétraitements afin de mettre en avant les parties en mouvements.

6.3 Estimation du mouvement via flot optique et différence d'images successives

Dans cette section le but est de ne plus traiter les images brutes dans lesquelles beaucoup d'informations inutiles pour notre application sont présentes mais d'appliquer certaines transformations afin que les vidéos mettent en avant de manière plus explicite les mouvements. Pour ça on va utiliser deux techniques : la différence d'images successives et le flot optique. Ensuite les vidéos transformées seront traitées de 2 façons, tout d'abord en ré-utilisant l'approche C3D et LSTM puis en créant nous-même les descripteurs pour expliquer le mouvement dans une séquence.

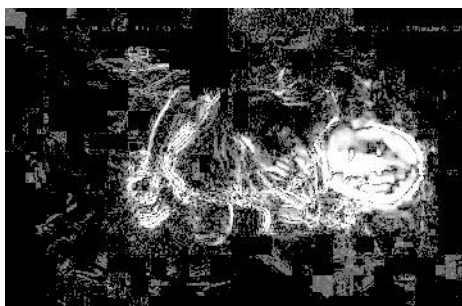
6.3.1 Extraction des informations de mouvements

Différences d'images successives On commence d'abord par effectuer la différence absolue de chaque image avec la suivante dans une vidéo. On définit $d(X_i, X_{i+1})$ la différence absolue entre deux images RGB X_i et X_{i+1} .

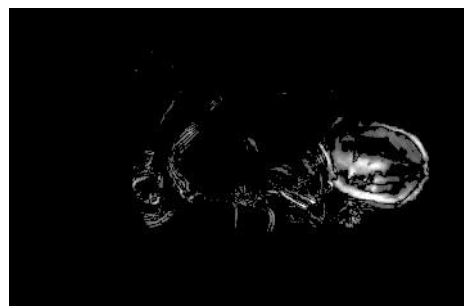
$$d(X_i, X_{i+1}) = \left| \frac{x_{i,1} + x_{i,2} + x_{i,3}}{3} - \frac{x_{i+1,1} + x_{i+1,2} + x_{i+1,3}}{3} \right|$$

Où $x_{i,1}$, $x_{i,2}$ et $x_{i,3}$ représentent les composantes R, G, B de l'image X_i .

Afin d'éviter de visualiser les changements trop légers qui peuvent apparaître dans les vidéos comme de légers changements de luminosité ou des déplacements de draps sur les matelas des enfants on applique un seuil qui conserve seulement les grandes variations de niveaux de gris. En appliquant cette transformation on obtient des images comme dans la Fig. 19.



(a)



(b)

Figure 19 - Extraits des vidéos avec différences absolues. On peut visualiser les bras et jambes de l'enfant (a) Différences d'images successives sans seuillage (exposition augmentée pour une meilleure visualisation) (b) Différences d'images successives avec seuillage.

Cette technique permet d'avoir une bonne visualisation de la localisation des déplacements mais on aimerait avoir une meilleure estimation en ce qui concerne la norme du déplacement et l'angle. Pour ça on va utiliser le flot optique.

Flot optique Un bon moyen d'estimer ces paramètres est le flot optique. Ici on utilise l'algorithme de Farneback [25] qui permet de calculer le flot optique pour tous les pixels de la vidéo. On obtient les résultats représentés dans la figure 20. Cette visualisation est obtenue en affichant l'angle et la norme dans une image en HSV (pour Teinte Saturation Valeur) en associant les angles à la teinte et la norme à la valeur.



Figure 20 - Extraits des vidéos avec flot optique. On observe un rendu plus nuancé avec plus d'informations que la différence absolue mais on perd en détail.

Différence absolue et flot optique Pour utiliser les avantages des deux méthodes on choisit d'utiliser notre rendu des différences absolues comme un masque binaire sur le flot optique. On perd la nuance qu'apportaient les différences absolues en les considérant comme binaire mais le flot optique pallie à ce problème. On obtient une nouvelle représentation comme représenté figure 21.

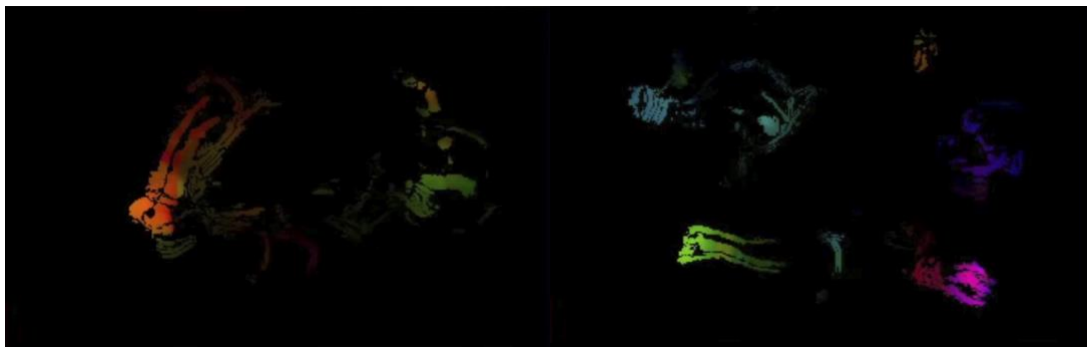


Figure 21 - Combinaison du flot optique et des différences d'images successives.

Dans un premier temps, ces vidéos pré-traitées sont utilisées avec le modèle C3D et les LSTMs.

6.3.2 C3D et LSTM

En reprenant le premier modèle, aucun résultat significatif n'a pu être obtenu. Ceci peut laisser à penser que le modèle C3D n'est pas adapté à notre application. On tente

donc, dans un dernier test, de le supprimer et d'extraire manuellement des caractéristiques du flot optique.

6.3.3 Création de descripteurs et LSTM

On souhaite donc, grâce aux informations de mouvements, créer un descripteur qui pourrait être mis en entrée des LSTMs. C'est à dire qu'on veut décrire chaque vidéo par une suite de vecteurs. La norme, l'angle et le nombre de pixels en mouvement (avec une valeur de différences successives dépassant le seuil) et surtout la variation de ces valeurs pourraient être de bons indicateurs pour décrire la fluidité voire la complexité du mouvement.

Chaque image est donc décrite par un vecteur représentant le mouvement de dimension 17. Tout d'abord on exprime le nombre de pixels en mouvement sous forme d'un ratio du nombre de pixels actifs sur le nombre de pixels de l'image. Ceci permet d'être invariant aux changements d'échelles. Pour l'angle et la norme du flot optique, on choisit de faire 2 histogrammes découpés en 8 sections chacun. On concatène ces descripteurs au cours du temps pour former une matrice dont un exemple peut-être visualisé figure 22.

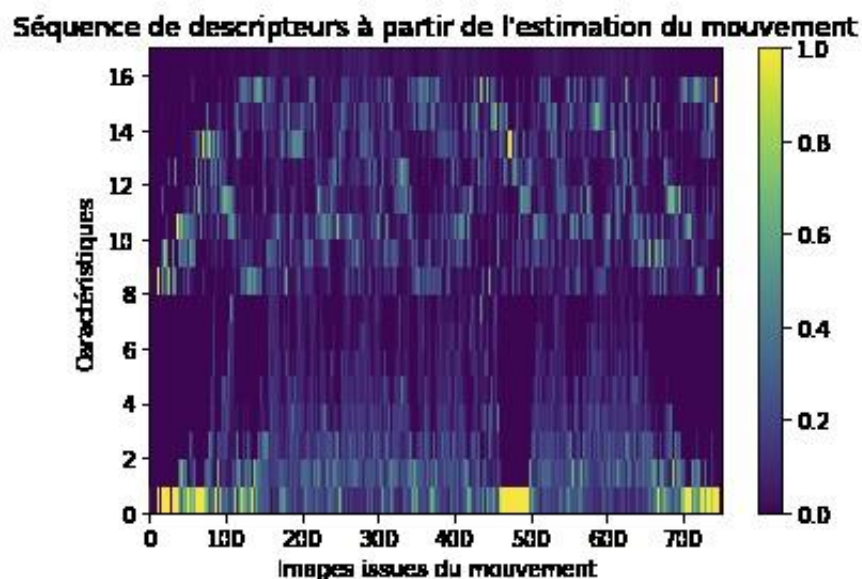


Figure 22 - Descripteur extrait du flot optique et des différences d'images successives.

Ensuite on lance un entraînement sur ces données avec un modèle LSTM. Après plusieurs essais on obtient le résultat de la figure 23 avec une matrice de confusion représentée dans la figure 24. On constate un sur-apprentissage quand on regarde la fonction perte mais surtout le taux de reconnaissance qui n'apporte pas de meilleurs résultats que le hasard, de plus la matrice de confusion sur l'ensemble de validation montre que le réseau n'apprend pas correctement.

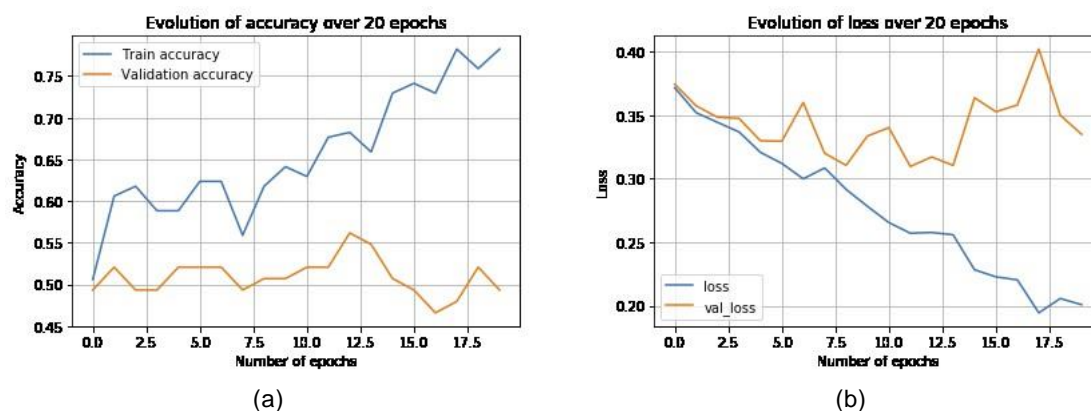


Figure 23 - Entraînement sur un modèle LSTM sur l'analyse du mouvement. (a) Graphique représentant l'évolution du taux de reconnaissance du modèle (b) Graphique représentant la fonction perte du modèle.

	Normal prédit	Anormal prédit
Normal véritable	27	9
Anormal véritable	27	9

Figure 24 - Matrice de confusion du modèle LSTM sur l'analyse du mouvement.

6.3.4 Conclusion

Les premiers résultats ne sont pas concluants. On peut en déduire que les descripteurs sont incapables de coder proprement les mouvements de l'enfant, leur complexité, variations et répétitivité. La difficulté vient du fait qu'on ne peut pas transmettre les mouvements liés aux membres et aux extrémités. S'il y a beaucoup de mouvement on ne sait pas si cela vient de la tête ou des bras par exemple.

6.4 Résultats obtenus et comparaison avec les objectifs

Dans cette partie nous avons développé trois principes différents afin de résoudre le problème de l'analyse des mouvements généraux spontanés. Premièrement nous avons utilisé le modèle C3D afin de décrire des courtes séquences d'actions et avons ensuite entraîné un modèle LSTM afin de prendre en compte les dépendances à long terme. Nous avons utilisé l'analyse en composante principale afin de conserver les informations de variance maximale dans les descripteurs. Dans un deuxième temps nous avons essayé une alternative aux modèles LSTMs avec les TCN appliqués aux caractéristiques C3D. Enfin nous avons analysé le mouvement de chacune des séquences grâce au flot optique et aux différences absolues.

Malgré les différents modèles mis en place nous n'avons pas constaté de prédictions suffisamment intéressantes pouvant mener à une appréciation globale des mouvements spontanés. Les objectifs n'ont donc pas été atteints, néanmoins nous avons introduits plusieurs modèles qui gardent un potentiel intéressant pour la description des GMs et émis de nouvelles idées afin d'aborder le problème différemment.

6.5 Difficultés rencontrées

Dans ce paragraphe seront avancées les difficultés rencontrées pendant le stage en parlant par exemple des éléments non mis en avant dans le rapport. Le stage ayant pour objet le développement d'un outil informatique faisant appel à des algorithmes d'apprentissage profond, il était nécessaire d'apprendre le fonctionnement de ces algorithmes mais aussi à les utiliser. Une grande partie du stage a donc consisté en une introduction à l'apprentissage profond. Dans un second temps une portion importante du temps du stage a consisté à mettre en forme les données c'est à dire les phases de pré-traitement mais aussi le passage des données brutes à des données annotées capable d'être utilisées. Toutes ces étapes n'étant pas l'objet du stage elles n'ont pas vraiment été abordées dans ce rapport. De plus toutes ces phases ont nécessité d'apprendre de nouveaux outils pour le traitement d'images, de vidéos mais aussi de traitements de fichiers textes par exemple. C'est une partie assez intéressante car j'ai dû développer des algorithmes qui m'ont permis d'automatiser certaines actions comme l'analyse du framerate des images ou la lecture des annotations dans un fichier excel.

De plus nous avons rencontrés des problèmes dans les données envoyées et il a donc été nécessaire de faire modifier l'ensemble de données. Cela mène à une autre difficulté. Les données traitées faisant appel au champ médical et leur analyse demandant des compétences particulières il était difficile de comprendre exactement ce qui différenciait une séquence normale d'une séquence anormale en regardant les vidéos.

Une autre difficulté était les entraînements d'apprentissage profond qui sont des étapes qui peuvent être longues. De plus on ajuste certains paramètres pour voir l'influence de celui-ci étant donné qu'il n'y a pas de méthodes à appliquer pour atteindre un résultat optimal. Ceci augmente le temps passé à faire des entraînements. Et au vu des résultats qui n'ont jamais dépassé le hasard il a fallu de nombreux entraînements infructueux.

7 Conclusion

Pour rappel, l'objectif de ce stage est de développer un outil informatique capable d'évaluer la qualité des GMs de nourrissons prématurés à haut-risque. On souhaite développer cet outil grâce à des algorithmes d'apprentissage profond qui doivent apprendre les caractéristiques des GMs normaux et anormaux via des séquences vidéos annotées. Après plusieurs essais les travaux effectués durant le stage n'ont pas permis d'obtenir des évaluations suffisamment significatives pour que l'outil soit efficace.

Le travail effectué pendant ce stage a consisté à appliquer trois approches différentes au même problème. Dans un premier temps nous avons utilisé un modèle C3D déjà entraîné sur les séquences vidéos et avons ensuite entraîné un modèle LSTM sur ces descripteurs. Ensuite nous avons utilisé un modèle d'encodeur TCN et enfin nous avons analysé les vidéos grâce à différentes techniques avant d'apprendre leurs caractéristiques.

Dans les 3 cas les résultats obtenus n'étaient pas supérieurs à une prédiction au hasard.

L'échec de l'objectif peut-être expliqué par plusieurs raisons. Tout d'abord les données traitées sont assez complexes et il est difficile pour une personne non entraînée de classifier les séquences en anormale ou normale. On observe aussi beaucoup de redondances dans les données, ce qui a pu causer nos modèles à subir un sur-apprentissage et dans le cas où l'on évite les séquences redondantes on obtient un petit ensemble de données, peut-être trop petit pour un apprentissage. En ce qui concerne C3D nous avons émis l'hypothèse qu'il était suffisant pour décrire les mouvements de l'enfant ce dont nous ne sommes pas certains. Enfin, par manque de temps, nous avons préféré, face aux mauvais résultats, essayer plusieurs approches sans aller au bout de ces méthodes, il n'est donc pas impossible que sous d'autres conditions ces méthodes soient efficaces, par exemple un ensemble de données plus grand ou décrit plus précisément.

Dans le document décrivant l'étude observationnelle il est fait mention la possibilité d'utiliser des algorithmes d'apprentissage non supervisé afin de classifier les séquences sans annotations, il pourrait être intéressant d'explorer cette piste.

Une dernière piste, qui semble porteuse, est d'annoter les instants où des mouvements anormaux apparaissent et de chercher à détecter ces instants. En effet, rechercher quels mouvements caractérisent les GMs anormaux pourrait permettre de simplifier le problème.

Références

- [1] M. Hadders-Algra, "Motricité spontanée normale et pathologique du jeune nourrisson," *Enfance*, vol. 55, no. 1, pp. 13-22, 2003.
- [2] H. F. R. Prechtl, "Qualitative changes of spontaneous movements in fetus and pre-term infant are a marker of neurological dysfunction," *Early Human Development*, vol. 23, no. 3, pp. 151-158, 1990.
- [3] F. Ferrari, G. Cioni, and H. F. Prechtl, "Qualitative changes of general movements in preterm infants with brain lesions," *Early Human Development*, vol. 23, no. 3, pp. 193-231, 1990.
- [4] C. Einspieler and H. F. Prechtl, "Prechtl's assessment of general movements : A diagnostic tool for the functional assessment of the young nervous system," *Mental Retardation and Developmental Disabilities Research Reviews banner*, vol. 11, no. 1, pp. 61-67, 2005.
- [5] K. P. Murphy, *Machine Learning : A Probabilistic Perspective*. Cambridge, MA : MIT Press, 2016.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, 2015.
- [7] Stanford Vision and Learning Lab | Stanford University, "Neural Networks Part 1 : Setting Up the Architecture," *Convolutional Neural Networks for Visual Recognition*. [Online]. Available : <http://cs231n.github.io/>[Accessed:11-June-2019]
- [8] Mayranna [CC BY-SA 3.0 (<https://creativecommons.org/licenses/by-sa/3.0/>)], "Perceptron," 2013, [Accessed July 10, 2019]. [Online]. Available : https://commons.wikimedia.org/wiki/File:Perceptron_moj.png
- [9] K. Fukushima, "Neocognitron : A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, pp. 193-202, 1980.
- [10] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, Ed. Morgan-Kaufmann, 1990, pp. 396-404. [Online]. Available : <http://papers.nips.cc/paper/293-handwritten-digit-recognition-with-a-back-propagation-network.pdf>
- [11] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov 1998.
- [12] Stanford Vision and Learning Lab | Stanford University, "Convolutional Neural Networks (CNNs / ConvNets)," *Convolutional Neural Networks for Visual Recognition*. [Online]. Available : <http://cs231n.github.io/>[Accessed:11-July-2019]
- [13] Aphex34 - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=45661858>, "Cnn layers arranged in 3-d volumes," 2015, [Accessed July 11, 2019]. [Online]. Available : https://commons.wikimedia.org/wiki/File:Conv_layers.png#/media/File:Conv_layers.png
- [14] S. Ji, M. Y. Wei Xu, and K. Yu, "3d convolutional neural networks for human action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 221-231, 2013.

- [15] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in CVPR, 2014.
- [16] D. Tran, L. D. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "C3D : generic features for video analysis," CoRR, vol. abs/1412.0767, 2014. [Online]. Available : <http://arxiv.org/abs/1412.0767>
- [17] Z. C. Lipton, "A critical review of recurrent neural networks for sequence learning," CoRR, vol. abs/1506.00019, 2015. [Online]. Available : <http://arxiv.org/abs/1506.00019>
- [18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
- [19] François Deloche [CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>)], "A diagram for a one-unit long short-term memory (Istm)," 2017, [Accessed August 28, 2019]. [Online]. Available : <https://commons.wikimedia.org/w/index.php?curid=60149410>
- [20] Y. Fan, X. Lu, D. Li, and Y. Liu, "Video-based emotion recognition using cnn-rnn and c3d hybrid networks," in Proceedings of the 18th ACM International Conference on Multimodal Interaction, ser. ICMI '16. New York, NY, USA : ACM, 2016, pp. 445-450. [Online]. Available : <http://doi.acm.org/10.1145/2993148.2997632>
- [21] V. Escorcia, F. Caba Heilbron, J. C. Niebles, and B. Ghanem, "Daps : Deep action proposals for action understanding," in Computer Vision - ECCV 2016, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham : Springer International Publishing, 2016, pp. 768-784.

- [22] L. Zhang, G. Zhu, P. Shen, J. Song, S. Afaq Shah, and M. Bennamoun, "Learning spatiotemporal features using 3dcnn and convolutional lstm for gesture recognition," in The IEEE International Conference on Computer Vision (ICCV) Workshops, Oct 2017.
- [23] R. Reed and R. Marks, Neural Smithing : Supervised Learning in Feedforward Artificial Neural Networks, ser. Bradford book. MIT Press, 1999. [Online]. Available : <https://books.google.fr/books?id=cnbgoQEACAAJ>
- [24] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," CoRR, vol. abs/1611.05267, 2016. [Online]. Available : <http://arxiv.org/abs/1611.05267>
- [25] G. Farnebäck, "Two-frame motion estimation based on polynomial expansion," In : Image analysis, vol. 2749, pp. 363-370, 06 2003.

