



**UNIVERSITÉ
TOULOUSE III**
PAUL SABATIER



UNIVERSITE PAUL SABATIER – TOULOUSE III
Faculté de médecine Toulouse Rangueil
Enseignement des techniques de réadaptation

MÉMOIRE

*Présenté en vue de l'obtention du
Certificat de Capacité d'Orthophonie*

**Etude de validité de critère d'une grille d'évaluation du
nourrisson lors de la prise alimentaire au biberon**

Par Charlotte STOSSKOPF

Sous la direction de :

Dr Pascale FICHAUX-BOURIN, médecin phoniatre

Mme Gabrielle DE MONTEIL, orthophoniste

Membres du jury :

Dr Pascale FICHAUX-BOURIN, médecin phoniatre

M. Christophe RIVES, orthophoniste

Mme Anaïs BARRY, orthophoniste

Juin 2022

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce mémoire, et en particulier :

Pascale Fichaux-Bourin et Gabrielle De Monteil pour m'avoir proposé ce sujet de mémoire et l'avoir encadré tout au long de l'année. Merci à toutes les deux pour leur soutien, leur confiance, leurs conseils et le temps qu'elles m'ont accordé.

L'équipe de recherche du centre de référence du syndrome de Prader-Willi de Toulouse, et notamment Catherine Molinas, pour leur disponibilité et la mise à disposition des données et informations nécessaires à la rédaction de ce mémoire.

Mathieu Balaguer pour sa disponibilité et son aide au niveau des analyses statistiques.

Les orthophonistes passionnés que j'ai rencontrés, pour leur savoir faire, leur accueil et leur volonté de transmettre.

Ma promotion et l'équipe pédagogique du Centre de formation, pour l'entraide et la bienveillance qui ont accompagné ces cinq dernières années riches en expériences et apprentissages.

Je remercie profondément l'ensemble de mes proches, famille et amis.

Mes parents et frères et sœurs pour leur soutien constant. Plus particulièrement mes sœurs, l'une pour sa présence et ses encouragements confiants, l'autre pour ses relectures et plannings tyranniques dans la dernière ligne droite, sans lesquels mon mémoire aurait beaucoup perdu en qualité et en rapidité de rédaction.

Les 12 sans qui ces cinq années n'auraient été aussi belles, enrichissantes et joyeusement déjantées. Pour leur soutien infailible, quelles que soient les circonstances : partiels, covid, CCC, mémoire... Rien n'a résisté à l'entraide et à la bonne humeur.

A ma coloc' des deux dernières années avec laquelle tous les aléas du quotidien ont été surmontés avec succès, de la légère fuite d'eau aux soirées de déconnexion cérébrale.

Et enfin, à la plus fidèle des acolytes à qui je devrai toujours la moitié de mon diplôme, la totalité de ma santé mentale et tellement d'autres choses qui seraient trop nombreuses à énumérer ici.

Table des matières

Remerciements	1
Table des matières	2
Glossaire	5
Table des illustrations.....	6
1 Introduction.....	7
2 Partie théorique	9
2.1 Evaluation des troubles de l'alimentation et de la déglutition chez l'enfant	9
2.1.1 Généralités	9
2.1.2 Evaluation instrumentale	11
2.1.3 Evaluation clinique	12
2.2 Présentation de l'outil élaboré.....	14
2.2.1 Description et propriétés cliniques de l'outil	14
2.2.2 Propriétés psychométriques de l'outil	15
2.2.3 Nécessité de poursuivre la validation de l'outil en vérifiant sa validité de critère	15
2.3 Méthodologie utilisée dans les outils existants	16
2.3.1 Revue des outils existants	16
2.3.2 Validité de critère : quelles méthodologies employées dans ces tests.....	18
2.4 Problématique et hypothèses	19
2.4.1 Problématique	19
2.4.2 Hypothèses.....	19
3 Méthodologie	20
3.1 Participants.....	20
3.2 Matériel et procédures.....	21
3.2.1 Procédure pour l'évaluation clinique	21
3.2.1.1 Présentation de la grille d'observation clinique.....	21
3.2.1.2 Protocole d'enregistrement vidéo	22
3.2.2 Procédure pour la vidéoradioscopie de la déglutition	23
3.2.2.1 Présentation de la grille de vidéoradioscopie	23
3.2.2.2 Protocole d'enregistrement vidéoradioscopique	24
3.2.3 Appariement des items des deux grilles	25
3.2.3.1 Item 1 : Continence labiale	25
3.2.3.2 Item 2 : Mouvements mandibulaires.....	25
3.2.3.3 Item 3 : Rythme de succion.....	26
3.2.3.4 Item 4 : Mouvements linguaux	26

3.2.3.5	Item 5 : Initiation de la déglutition pharyngée.....	26
3.2.3.6	Item 6 : Propulsion pharyngée.....	26
3.2.3.7	Item 7 : Continence vélopharyngée.....	27
3.2.3.8	Item 8 : Protection des voies respiratoires.....	27
3.2.3.9	Appariement des systèmes de notation.....	27
3.2.4	Recueil et préparation des données.....	28
3.3	Analyse des données - Statistiques.....	28
3.3.1	Démarche générale.....	28
3.3.2	Analyse des scores globaux.....	29
3.3.3	Analyse des données item par item.....	29
3.3.4	Les différents paramètres étudiés.....	30
3.3.4.1	L'aire sous la courbe (AUC).....	30
3.3.4.2	Les qualités intrinsèques.....	30
3.3.4.3	Les qualités extrinsèques.....	32
4	Résultats.....	33
4.1	Analyse des scores globaux.....	33
4.1.1	Résultats du protocole A.....	33
4.1.1.1	Statistiques descriptives.....	33
4.1.1.2	Courbe ROC.....	34
4.1.2	Résultats du protocole B.....	36
4.1.2.1	Statistiques descriptives.....	36
4.1.2.2	Courbes ROC.....	37
4.1.3	Comparaison des deux protocoles.....	38
4.2	Analyse des résultats item par item.....	40
4.2.1	Résultats du protocole A.....	40
4.2.2	Résultats du protocole B.....	42
4.2.3	Comparaison des deux protocoles.....	44
5	Discussion.....	46
5.1	Vérification des hypothèses.....	46
5.1.1	Hypothèse 1 : L'outil testé possède dans son ensemble une bonne validité de critère.....	46
5.1.2	Hypothèses 2 : Les items de l'outil testé possèdent une bonne validité de critère.....	46
5.2	Limites et biais.....	46
5.2.1	Choix de la population.....	46
5.2.2	Procédures de passation et cotation.....	47
5.2.3	Sélection et appariement des items.....	48
5.2.4	Choix du Gold Standard.....	49

5.3	Retour d'expérience et applications futures.....	49
5.3.1	Retour sur les résultats des scores globaux	49
5.3.2	Retour sur les résultats par item	51
5.3.3	Ebauche de validation de la sensibilité au changement	52
5.4	Perspectives.....	53
6	Conclusion	54
	Bibliographie.....	55
	Table des annexes	60
	Annexes	61

GLOSSAIRE

AUC : Area Under Curb

BBAT : Bristol Breastfeeding Assessment Tool

COSMIN : COnsensus-based Standards for the selection health Measurement Instruments

Courbes ROC : Receiver Operating Characteristic

EFS : the Early Feeding Skills Assessment

FN : Faux Négatifs

FP : Faux Positifs

IBFAT : Infant Breastfeeding Assessment Tool

J : Indice de Youden

MBA : the Mother-Baby Assessment

NNS : Non-Nutritive sucking Scoring System

NOMAS : Neonatal Oral-Motor Assessment Scale

OFS : Oral Feeding Skills

PIBBS : the Preterm Infant Breastfeeding Behaviour Scale

SAIB : Systematic Assessment of the Infant at Breast

Se : Sensibilité

Spe : Spécificité

SPW : Syndrome de Prader-Willi

VN : Vrais Négatifs

VP : Vrais Positifs

VPN : Valeur Prédictive Négative

VPP : Valeur Prédictive Positive

VRDG : VidéoRadioscopie de la DéGlutition

TABLE DES ILLUSTRATIONS

Tableau 1 : Aperçu de la grille d'observation clinique utilisée pour les cotations de tétée	22
Tableau 2 : Aperçu de la grille utilisée pour la cotation vidéoradioscopique	24
Tableau 3 : Aperçu du tableau d'appariement pour l'analyse globale et par item - Protocoles A et B	28
Tableau 4 : Tableau de contingence entre la grille d'observation et la grille de VRDG	30
Tableau 5 : Tableau de contingence pour la valeur seuil de 3/8 - Protocole A.....	35
Tableau 6 : Tableau de contingence pour la valeur seuil de 1/8 - Protocole B.....	37
Tableau 7 : Sensibilité (Se), Spécificité (Spe) et indice de Youden (J) de chaque item - Protocole A ...	41
Tableau 8 : Sensibilité (Se), Spécificité (Spe) et indice de Youden (J) de chaque item - Protocole B ...	43
Tableau 9 : Représentation concrète des 32 vidéos en fonction de leur statut normal/anormal et de leur résultat à la grille d'observation - Protocole A	50
Tableau 10 : Représentation concrète des 32 vidéos en fonction de leur statut normal/anormal et de leur résultat à la grille d'observation - Protocole B.....	50
Figure 1 : Notation de chaque vidéo via la grille d'observation clinique - Protocole A	33
Figure 2 : Courbe ROC obtenue pour les scores de la grille d'observation clinique - Protocole A	34
Figure 3 : Notation de chaque vidéo via la grille d'observation clinique - Protocole B	36
Figure 4 : Courbe ROC obtenue pour les scores de la grille d'observation clinique - Protocole B.....	37
Figure 5 : Comparaison des répartitions du nombre de Vrais Positifs, Vrais Négatifs, Faux Positifs et Faux Négatifs pour chaque protocole	39
Figure 6 : Comparaison des valeurs de la Sensibilité et de la Spécificité des deux protocoles	39
Figure 7 : Comparaison des valeurs de la VPP et de la VPN entre les deux protocoles.....	40
Figure 8 : Répartition par item du nombre de VP, VN, FP et FN - Protocole A	40
Figure 9 : Représentation de la Sensibilité (Se), de la Spécificité (Spe) et de l'indice de Youden (J) pour chaque item - Protocole A.....	42
Figure 10 : Répartition par item du nombre de VP, VN, FP et FN - Protocole B	42
Figure 11 : Représentation de la Sensibilité (Se), de la Spécificité (Spe) et de l'indice de Youden (J) pour chaque item - Protocole B	44
Figure 12 : Comparaison de la Sensibilité et de la Spécificité par item selon les protocoles A et B.....	44
Figure 13 : Comparaison des indices de Youden pour chaque item selon les protocoles A et B	45

1 INTRODUCTION

Les troubles de l'alimentation et de la déglutition dans la population pédiatrique présentent une incidence élevée qui tend à augmenter avec le temps (Arvedson, 2008; Dodrill & Gosa, 2015; Lefton-Greif, 2008). Ceci peut s'expliquer par les avancées scientifiques et médicales qui permettent un meilleur repérage et un diagnostic plus précoce de ces troubles, ainsi que de meilleures chances de survie chez les enfants à risque (Lefton-Greif, 2008). Des difficultés d'alimentation et de déglutition trop importantes ou qui se chronicisent risquent fortement d'impacter le développement de l'enfant et sa qualité de vie à long terme (Cascales, 2015). Il est donc essentiel de pouvoir repérer et diagnostiquer le plus précocement possible les enfants à risque, ainsi qu'identifier les mécanismes physiopathologiques impliqués, d'où la nécessité d'avoir à disposition des outils d'évaluation fiables et adaptés.

Les évaluations instrumentales, souvent considérées comme examens de référence, ont cependant besoin d'être complétées par des outils cliniques, en raison du caractère invasif et non écologique des évaluations instrumentales qui doit rendre leur utilisation la plus ponctuelle et contrôlée possible (Arvedson & Lefton-Greif, 2017). Cependant, il existe peu d'outils permettant l'évaluation clinique d'enfants alimentés au biberon. De plus, parmi ces outils, tous ne correspondent pas à l'âge et à la pathologie de l'enfant ou ne possèdent pas des propriétés psychométriques satisfaisantes (Bickell et al., 2017).

C'est en partant de ce constat que Morgane Péron a élaboré en 2018 un outil d'évaluation non-instrumentale des capacités d'alimentation et de déglutition du nourrisson alimenté au biberon, dans le cadre de son mémoire en vue de l'obtention du certificat de capacité d'orthophonie. Celui-ci faisait suite à l'étude préliminaire de Pauline Crusson sur ce sujet, réalisée en 2013 (Crusson, 2013).

Cette grille s'adresse aux professionnels de soins de la petite enfance, et correspond à tout nourrisson de 0 à 6 mois, alimenté au biberon. Elle a pour objectif de permettre une orientation plus rapide et pertinente de l'enfant vers des examens complémentaires ou prises en charge précoces des éventuels troubles d'alimentation ou de déglutition (Péron, 2018).

La standardisation d'un test tel que celui-ci dépend de la validation de ses qualités psychométriques. Celles-ci comprennent différents paramètres dont les volets principaux sont la validité, la fidélité et la sensibilité au changement (Fermanian, 2005; Mokkink et al., 2010).

Validité de contenu, fidélité test-retest, et fidélité intra-juge ont pu être vérifiées par les précédents travaux sur le sujet (Péron, 2018; Poumarède, 2019).

En outre, le mémoire actuel d'Emilie Leboucher a tenté de valider la fidélité inter-juges en tenant compte des limitations relevées dans les travaux précédents.

En parallèle, l'objectif de ce présent travail était de poursuivre la validation de cet outil d'évaluation non-instrumentale en démontrant sa validité de critère, de manière quantitative cette fois. Une cotation standardisée de la vidéoradioscopie de la déglutition à corrélérer avec la grille d'observation a été utilisée.

Dans une première partie théorique, nous décrivons les principaux moyens d'évaluation des troubles de l'alimentation et de la déglutition chez le nourrisson, nous présentons l'outil élaboré, et nous expliquons la méthodologie employée pour la validation des outils existants.

Nous exposons par la suite la problématique et précisons la méthodologie utilisée lors de cette étude.

Enfin, nous détaillons les résultats obtenus et nous discutons de ces derniers et des possibilités futures d'amélioration de la grille.

2 PARTIE THEORIQUE

2.1 EVALUATION DES TROUBLES DE L'ALIMENTATION ET DE LA DEGLUTITION CHEZ L'ENFANT

2.1.1 Généralités

Les troubles de l'alimentation et de la déglutition regroupent une grande diversité de difficultés pouvant apparaître autour des temps de repas. Celles-ci peuvent être d'origine anatomique ou fonctionnelle. Dans la population pédiatrique, elles peuvent aussi bien concerner les enfants atteints de troubles développementaux que ceux sans antécédent ou problème de santé particulier (Duffy, 2018).

Il est cependant important de bien différencier troubles de l'alimentation et troubles de la déglutition qui représentent deux problématiques bien différentes. En effet, un enfant peut avoir des troubles alimentaires sans nécessairement qu'il y ait de troubles de la déglutition, et inversement (Arvedson, 2008).

Le trouble alimentaire se caractérise par des difficultés pouvant toucher différents aspects de l'acte d'alimentation. Face au manque de consensus quant à la définition de ces troubles dans le domaine pédiatrique, Goday et al. (2019) a récemment proposé de les regrouper sous l'appellation « Troubles Alimentaires Pédiatriques » (Pediatric Feeding Disorders). Il regroupe ainsi sous ce terme tout trouble correspondant à cette définition : « une altération de la prise alimentaire orale qui n'est pas adaptée à l'âge et qui est associée à des difficultés médicales, nutritionnelles, des compétences alimentaires et/ou à un dysfonctionnement psychosocial. » [Notre Traduction].

Le trouble de la déglutition, anciennement nommé « dysphagie », concerne quant à lui toute perturbation d'une ou plusieurs phases de la déglutition, lesquelles sont la phase de préparation orale, la phase orale, la phase pharyngée et la phase œsophagienne. Il est diagnostiqué par le biais d'une évaluation clinique ainsi que d'une évaluation instrumentale (Arvedson, 2008; Duffy, 2018). Chez l'enfant, on considère que ce type de trouble peut impacter aussi bien les phases de la déglutition, l'interface entre déglutition et respiration, ou les deux (Arvedson & Lefton-Greif, 2017).

Lawlor & Choi (2020) décrivent, comme principaux symptômes de ces troubles, l'allongement du temps de repas, la diminution de l'appétence et le refus alimentaires, le ralentissement du développement, le reflux nasal, la toux, la respiration mouillée pendant et après le repas, et la

respiration laborieuse. Ils rappellent également la possible survenue de pénétrations et/ou inhalations laryngées, ces dernières pouvant s'avérer silencieuses.

Ces troubles peuvent avoir une étiologie congénitale ou bien acquise. La sévérité et l'évolution dépendent de nombreux facteurs tels que les troubles associés, les habitudes alimentaires, l'état de santé et le développement global de l'enfant (Arvedson & Lefton-Greif, 2017).

Concernant l'incidence de ces troubles de l'alimentation et de la déglutition, les chiffres sont incertains et varient beaucoup selon les critères diagnostiques de chaque professionnel. Tous s'accordent néanmoins à dire que leur fréquence a tendance à augmenter avec le temps, notamment chez les enfants prématurés ou présentant des comorbidités (Arvedson, 2008; Lawlor & Choi, 2020).

Cette évolution peut s'expliquer de différentes manières. D'une part, par la meilleure reconnaissance de ces troubles dans la population pédiatrique. D'autre part, du fait des progrès médicaux, chirurgicaux et technologiques permettant un meilleur taux de survie des enfants prématurés et/ou au contexte médical complexe (Arvedson & Lefton-Greif, 2017; Lefton-Greif & Arvedson, 2016).

Sur ces dernières années, Lawlor & Choi rapportent ainsi que 50% de parents d'enfants au développement typique constatent des troubles alimentaires, contre 80% chez des enfants présentant des troubles développementaux. Concernant les troubles de la déglutition, leur incidence serait de 0,9% environ dans la population générale, mais plus élevée dans les populations à risque (Lawlor & Choi, 2020).

L'évaluation de l'alimentation et de la déglutition a pour but d'identifier ces potentiels troubles chez l'enfant, de déterminer les causes et facteurs, et d'identifier les obstacles et les éléments facilitateurs au cours des temps de repas (Brodsky et al., 2019).

Elle doit être la plus précoce et exhaustive possible. Un travail pluridisciplinaire est donc essentiel entre professionnels de santé de la petite enfance et entourage familial, afin de permettre une vision globale de l'enfant et ainsi une prise de décision collaborative et un meilleur suivi (Arvedson, 2008; Kleinert, 2017).

Cette évaluation comprend un examen clinique, que nous détaillerons par la suite, et peut être complétée ou non par une évaluation instrumentale (Arvedson, 2008; Duffy, 2018).

2.1.2 Evaluation instrumentale

La réalisation d'une évaluation instrumentale chez l'enfant suspecté de présenter des troubles de l'alimentation et/ou de la déglutition n'est pas systématique. La décision de l'effectuer ou non dépend effectivement de plusieurs facteurs. Les professionnels de santé se questionnent notamment sur l'utilité et le probable impact que ces informations auront sur la pose de diagnostic et la prise en charge de l'enfant, comparé aux risques que certains examens peuvent avoir sur sa santé générale. Ils cherchent également à s'adapter à l'âge et aux capacités du patient (Arvedson & Lefton-Greif, 2017; Brodsky et al., 2019).

Les outils d'évaluation instrumentale les plus employés en pédiatrie sont la vidéoradioscopie de la déglutition (VRDG ou vidéofluoroscopie) et la nasofibroskopie de la déglutition. Plus rarement et selon les situations, on peut également avoir recours à l'échographie (ou ultrasonographie de la déglutition) ou à l'auscultation cervicale (Arvedson, 2008; Arvedson & Lefton-Greif, 2017).

Selon Lefton-Greif (2016), ces outils présentent à peu de choses près les mêmes avantages et inconvénients chez les enfants que chez les adultes.

L'objectif principal de la nasofibroskopie est de vérifier la symétrie de la mobilité laryngo-pharyngée ainsi que l'absence d'obstruction des voies aériennes supérieures. La VRDG, quant à elle, a plutôt pour visée de quantifier et qualifier les troubles de la déglutition. C'est donc à ce deuxième examen que nous allons nous intéresser dans cette partie.

La VRDG est l'examen de référence pour l'exploration fonctionnelle de la déglutition. Il permet d'obtenir des images dynamiques des différents temps de la déglutition (Arvedson, 2008) et ainsi d'en étudier l'anatomie et la physiologie, de même que les mécanismes physiopathologiques. A partir des conclusions tirées de cet examen, le patient pourra donc être orienté vers un traitement et une prise en soin appropriés (Arvedson & Lefton-Greif, 2017).

La VRDG présente néanmoins des limites.

Tout d'abord, il est important de prendre en considération l'aspect irradiant de cet outil d'évaluation. Il y a effectivement un effet cumulatif de l'exposition aux radiations, d'où l'importance de s'assurer de la nécessité d'effectuer un tel examen (Arvedson & Lefton-Greif, 2017).

D'autre part, cette méthode d'évaluation requiert la coopération du patient. Si celui-ci refuse l'aliment, aucune analyse de la déglutition ne sera possible (Lawlor & Choi, 2020)

Par ailleurs, le produit de contraste baryté utilisé n'est pas représentatif des goûts et textures alimentaires habituels et donc rarement révélateur d'une situation réelle d'alimentation. Il est également important de noter l'impossibilité d'évaluer l'alimentation au sein chez le nourrisson. De plus, cet examen ne permet d'observer qu'un petit nombre de déglutitions, à un instant T, dans des conditions non écologiques. Pour toutes ces raisons, il est important de ne pas interpréter de manière dichotomique normal/pathologique les résultats et observations obtenus. L'histoire et les informations globales du patient, les autres tests diagnostiques et les facteurs environnementaux sont à prendre en compte afin de cerner au mieux le profil et les difficultés du patient (Arvedson, 2008; Arvedson & Lefton-Greif, 2017; Brodsky et al., 2019; Duffy, 2018).

Chez l'enfant, la question de l'exposition aux radiations est d'autant plus préoccupante que ce dernier présente une plus forte sensibilité aux dommages cellulaires et des risques d'irradiations répétées sur toute une vie.

Les professionnels de santé se doivent donc de justifier une telle décision en évaluant le rapport bénéfices-risques. Pour cela, 4 facteurs principaux sont à considérer : existence d'une suspicion de dysphagie oropharyngée, attente d'une clarification du diagnostic et d'une aide aux décisions, bonnes dispositions de l'enfant pour y participer, et probabilité pour les résultats de marquer une différence dans la prise en soin de l'enfant.

Une irradiation minimale sera également privilégiée. Cela dépendra de l'équipement à disposition, des procédures, du personnel, et des facteurs spécifiques à l'enfant (Arvedson & Lefton-Greif, 2017).

2.1.3 Evaluation clinique

L'évaluation non-instrumentale est une étape primordiale dans la recherche de troubles de l'alimentation et de la déglutition chez l'enfant, ainsi que dans l'aide à la décision thérapeutique. Elle doit être réalisée systématiquement, avant toute prise de décision d'évaluation instrumentale (Lefton-Greif & Arvedson, 2016).

Les objectifs de cet examen sont d'identifier les potentiels troubles et de poser des hypothèses sur leur nature, sévérité et étiologies possibles.

Il permet une analyse anatomique des structures, principalement orales, et une analyse fonctionnelle des capacités orales sensori-motrices et respiratoires.

Le recueil de toutes ces informations aide ainsi les professionnels de santé à déterminer les différentes options alimentaires qui se présentent, les éléments qui peuvent bloquer ou au contraire aider à l'alimentation et à la déglutition du patient.

Enfin, au terme de cette évaluation se pose la question de l'examen instrumental. La question est de savoir s'il est nécessaire et envisageable compte tenu des risques et caractéristiques du patient. Dans ce dernier cas, on se questionne sur l'outil à utiliser. Il est essentiel pour le professionnel de santé de la petite enfance de réfléchir à ces différents aspects afin d'orienter au mieux la recherche diagnostique et la potentielle prise en soin (Arvedson, 2000; Brodsky et al., 2019; Lefton-Greif & Arvedson, 2016).

L'évaluation clinique comprend différentes étapes. Il s'agit tout d'abord de réaliser une anamnèse reprenant les éléments clés de l'histoire familiale, développementale, alimentaire et sur le plan médical de manière plus globale (Arvedson, 2000).

Vient ensuite l'examen physique. Il consiste en un bilan analytique sensori-moteur oral des structures anatomiques, au repos puis en mouvement. Les capacités fonctionnelles sont également évaluées au moyen d'un essai alimentaire, avec de l'eau et/ou des aliments (Arvedson, 2008; Lawlor & Choi, 2020).

Enfin, une évaluation alimentaire est réalisée, en situation la plus écologique possible. Les professionnels de santé procèdent à une observation de repas, dans les conditions habituelles d'alimentation du patient (Arvedson, 2008).

Cependant, s'il existe plusieurs outils normalisés et reconnus concernant l'évaluation instrumentale, ce n'est pas le cas du côté de l'évaluation clinique, notamment chez le nourrisson. Celle-ci est effectivement bien moins documentée, et si de nombreux tests et questionnaires existent, très peu possèdent des qualités psychométriques fiables et démontrées (Heckathorn et al., 2016).

Selon Lawlor et Choi (2020), les questionnaires de screening des troubles de déglutition qui existent à ce jour ne concernent que des populations spécifiques porteuses de telle ou telle pathologie, et non pas la population pédiatrique générale.

Heckathorn (2016) souligne ainsi le réel besoin en outils valides dans ce domaine et la nécessité de réaliser une analyse sérieuse des outils existants telle que décrite dans le consensus COSMIN (Mokkink et al., 2010) (ANNEXE 1).

2.2 PRESENTATION DE L'OUTIL ELABORE

2.2.1 Description et propriétés cliniques de l'outil

La grille d'évaluation de l'alimentation et de la déglutition ainsi élaborée par Morgane Péron en 2018 s'adresse à tout nourrisson entre 0 et 6 mois alimenté au biberon, quel que soit son trouble ou sa pathologie. Elle est à destination de tous les professionnels de santé de la petite enfance.

Elle s'inspire de plusieurs outils d'évaluation tels que l'OFS (Oral Feeding Skills), la NOMAS (Neonatal Oral-Motor Assessment Scale) et l'EFS (the Early Feeding Skills Assessment), mais entend s'intéresser plus spécifiquement aux mécanismes physiopathologiques et à leur sévérité.

Son objectif est de réaliser une évaluation rapide et plus systématique des capacités alimentaires du nourrisson. Cela permettrait ainsi de proposer des aides et adaptations alimentaires de manière plus précoce, et de réorienter rapidement les enfants les plus à risques vers les examens complémentaires appropriés. En effet, et comme vu plus haut, certaines de ces évaluations instrumentales ne sont pas sans danger pour les nourrissons et peuvent parfois n'être accessibles qu'après plusieurs mois d'attente, retardant ainsi la prise en soin et impactant le pronostic.

Cette grille peut également être intégrée au bilan orthophonique des troubles de la déglutition, de l'alimentation et de l'oralité chez le nourrisson.

Cinq parties composent la grille et s'intéressent aux différents temps de l'alimentation :

- Observation avant le repas, décrivant la préparation du repas et la disponibilité de l'enfant
- Observation des capacités motrices des organes impliqués, au temps oral plus spécifiquement
- Observation du déroulement de la séquence de déglutition, donnant des indices sur la qualité du temps pharyngé
- Observation du comportement global de l'enfant lors du repas pour avoir un premier aperçu des difficultés alimentaires autres que motrices
- Observation après repas, avec bilan alimentaire et recherches de signes indirects du déroulement du temps œsophagien

2.2.2 Propriétés psychométriques de l'outil

Dans son mémoire réalisé en 2018, Morgane Péron a pu commencer à vérifier les propriétés psychométriques de la grille élaborée.

La validité de contenu a été jugée satisfaisante par des professionnels de santé non spécialisés dans le domaine des troubles de l'alimentation et de la déglutition, potentiels futurs utilisateurs de la grille. Certaines modifications ont été apportées à l'outil initial et une grille plus courte a été créée avec des items vulgarisés, à destination des professionnels non-experts.

L'étude de la sensibilité a permis d'objectiver une différence significative entre un groupe de nourrissons présentant des troubles alimentaires et un groupe de nourrissons sans difficulté, validant par là-même ce critère.

Ce mémoire a également pu aborder la fidélité test-retest et la validité de critère. La première n'a pu être démontrée, soit du fait du délai d'une semaine trop long entre deux observations, soit du fait de performances trop variables d'une tétée à l'autre. La validité de critère, quant à elle, bien que présentant des résultats concordants, n'a pu être testée que sur le plan qualitatif, en comparant les résultats du test au compte-rendu de la VRDG pour sept nourrissons.

Un an plus tard, en 2019, Camille Poumarède s'est donc intéressée à la suite de la validation de cet outil. Les fidélités test-retest, intra-juge et inter-juges ont été revues. Cette dernière n'ayant pas été concluante, Emilie Leboucher reprend ce paramètre psychométrique dans son mémoire d'orthophonie réalisé au cours de cette année 2022.

2.2.3 Nécessité de poursuivre la validation de l'outil en vérifiant sa validité de critère

Les propriétés psychométriques ont été redéfinies il y a quelques années dans un article publié par l'initiative COSMIN (Consensus-based Standards for the selection health Measurement Instruments). Les spécialistes sont arrivés ensemble à un consensus à propos de la terminologie et des définitions des mesures psychométriques attendues lors de l'élaboration d'un test (Mokkink et al., 2010).

Les trois principaux domaines habituellement étudiés dans le cadre de la validation d'un outil sont ainsi : la fiabilité (reliability), la validité (validity) et la sensibilité au changement (responsiveness). Nous nous intéressons dans cette étude au volet validité. Celui-ci comprend la validité de contenu (content validity), la validité de construit (construct validity) et la validité de critère

(criterion validity). C'est à cette dernière que nous nous intéressons plus particulièrement dans cette étude.

La validité de critère (ou validité contre critère) est une propriété psychométrique qui vise à vérifier qu'un nouvel outil mesure fidèlement le phénomène qu'il est censé mesurer. Pour ce faire, il s'agit de comparer les résultats d'un même patient évalué par l'outil testé puis par un outil de référence (appelé gold standard). On considère que l'outil présente une bonne validité de critère si le gold standard est réellement considéré comme un outil de référence, et si les résultats de ces deux outils présentent une corrélation acceptable. Elle peut être étudiée de deux manières, selon les conditions de l'étude :

- la validité concourante ou concomitante (concurrent validity), à utiliser si les deux tests comparés sont administrés le même jour ou bien de manière rapprochée,
- la validité prédictive (predictive validity), à privilégier si plusieurs semaines ou mois se sont écoulés entre les deux tests (Fermanian, 2005; Mokkink et al., 2010; Terwee et al., 2007).

Comme précisé plus haut, Morgane Péron n'avait pu effectuer que la prévalidation de la validité de critère. Une comparaison qualitative avait été réalisée à partir des résultats objectivés provenant des compte-rendus de la VRDG. Si les résultats obtenus montraient une certaine concordance, cela ne permettait cependant pas de confirmer la validité de ce paramètre.

Comme relevé par Morgane Péron au terme de son travail, une vérification de la validité de critère sur le plan quantitatif est donc nécessaire. L'utilisation d'une échelle de cotation standardisée de la VRDG corrélée à la grille d'observation étudiée permettrait ainsi de confirmer la bonne validité de critère de cet outil.

2.3 METHODOLOGIE UTILISEE DANS LES OUTILS EXISTANTS

2.3.1 Revue des outils existants

Il existe un grand nombre d'outils cliniques élaborés depuis plusieurs années dans le domaine de l'alimentation et de la déglutition dans la population pédiatrique. Tous n'ont pas les mêmes objectifs et les mêmes propriétés cliniques et psychométriques. Afin de faire le tri parmi tous ces tests et d'y voir plus clair, plusieurs revues des outils cliniques existants ont été réalisées ces dernières années (Bickell et al., 2017; Heckathorn et al., 2016; Howe et al., 2008).

Ainsi, Bickell et al. (2017) a sélectionné les dix outils d'évaluation les plus pertinents et décrit leurs principales caractéristiques.

La NNS (Non-Nutritive sucking Scoring System) ne concerne, comme son nom l'indique, que la succion non-nutritive. Nous n'en parlerons donc pas dans ce travail.

Parmi les 9 outils restants, 3 concernent la prise au biberon :

- L'OFS (Oral Feeding Skills) : Ne s'intéresse qu'à la prise au biberon, évalue les capacités alimentaires orales, mais aucune étude des qualités psychométriques réalisée jusque-là.
- L'EFS (the Early Feeding Skills assessment) : S'intéresse à l'alimentation au biberon et au sein, évalue les capacités orales et d'alimentation et de déglutition, mais aucune étude des qualités psychométriques réalisée jusque-là.
- La NOMAS (the Neonatal Oral-Motor Assessment Scale) : S'intéresse à l'alimentation au biberon et au sein, évalue les capacités orales et d'alimentation et de déglutition, et plusieurs études évaluent ses propriétés cliniques et psychométriques (Howe et al., 2008).

Les 6 outils restants évaluent uniquement la prise alimentaire au sein :

- Aucune étude n'a encore évalué les propriétés psychométriques du SAIB (Systematic Assessment of the Infant at Breast).
- Le BBAT (Bristol Breastfeeding Assessment Tool) et le PIBBS (the Preterm Infant Breastfeeding Behaviour Scale) ont vu leurs qualités psychométriques analysées chacun par une étude.
- Les qualités psychométriques du LATCH, de l'IBFAT (Infant Breastfeeding Assessment Tool) et du MBA (the Mother-Baby Assessment) ont été évaluées par plusieurs études. Les résultats d'une étude à l'autre sont donc très variables.

2.3.2 Validité de critère : quelles méthodologies employées dans ces tests

Parmi ces dix outils relevés par Bickell et al. (2017), trois n'ont donc jamais vu leurs propriétés psychométriques testées (l'EFS, l'OFS et le SAIB), et deux n'ont pu vérifier que le principe de fiabilité (l'IBFAT et le MBA).

Les cinq autres, en revanche, sont à l'origine d'au moins une étude chacun vérifiant la fiabilité et la validité (le BBAT, le LATCH, la NOMAS, la NNS et le PIBBS). Parmi ces derniers, seuls le LATCH et la NOMAS sont concernés par des études de la validité de critère qui nous intéressent pour ce devoir.

Pour tester la validité de critère du LATCH, Kumar et al. (2006) a donc utilisé des courbes ROC (Receiver Operating Characteristic). Ces dernières sont une « représentation graphique de la relation existante entre sensibilité et spécificité d'un test, calculée pour toutes les valeurs seuils possibles » (Delacour et al., 2005). Elle permet ainsi de déterminer la valeur seuil optimale d'un test grâce à l'étude de l'aire sous la courbe (AUC : Air Under Curve). Après la fixation de ce seuil, Kumar et al. (2006) a donc calculé la spécificité et la sensibilité, ainsi que leurs intervalles de confiance. L'indice de Youden « J » a également été étudié. Celui-ci correspond à la somme des deux premières valeurs et donne ainsi une indication supplémentaire sur la précision de la méthode de diagnostic.

Concernant la NOMAS, Zarem et al. (2013) a étudié la validité concurrente de manière générale en effectuant des tests statistiques tels que celui de Student, de Mann-Whitney ou des régressions linéaires entre les scores de la NOMAS et plusieurs facteurs médicaux (mesures comportementales et des structures cérébrales de l'enfant). Les chercheurs se sont appuyés sur des enregistrements vidéo des temps d'alimentation des enfants pour le recueil de données.

En 2005, DeMatteo et al. a cherché à évaluer la précision de l'évaluation clinique des enfants présentant des difficultés d'alimentation et de déglutition, en particulier concernant les inhalations et pénétrations laryngées. Pour cela, un formulaire d'évaluation clinique des capacités oro-motrices et de la déglutition a été créé, à partir de l'évaluation clinique typique réalisée en milieu hospitalier et des données théoriques provenant de la littérature dans ce domaine. Ce formulaire a été testé par des professionnels de santé lors d'observations de repas d'enfants de 0 à 15 ans, dans des conditions les plus écologiques possibles (positions, aliments, ustensiles et aidants habituels). Les résultats ont ensuite été comparés à ceux de la VRDG de chacun de ces patients. Cette dernière servait donc d'outil de référence pour déterminer la validité de critère de ce test clinique en calculant la sensibilité, la spécificité et les valeurs prédictives positive et négative (DeMatteo et al., 2005).

2.4 PROBLEMATIQUE ET HYPOTHESES

2.4.1 Problématique

L'objectif de cette étude est de participer à la validation de l'outil d'évaluation non instrumentale élaboré au cours des travaux précédents.

Comme nous l'avons vu précédemment, certains critères psychométriques n'ont pas encore pu être démontrés. C'est le cas de la validité de critère (ou validité contre-critère), à laquelle nous allons donc nous intéresser dans cette étude.

Notre questionnement portera donc sur la capacité de cette grille d'évaluation clinique à détecter le plus fidèlement possible la présence ou l'absence de difficultés d'alimentation et de déglutition chez le nouveau-né nourri au biberon.

Notre travail actuel se concentrera plus particulièrement sur le volet déglutition, objectivable en VRDG, que sur le volet alimentation.

2.4.2 Hypothèses

Hypothèse 1 : L'outil testé possède dans son ensemble une bonne validité de critère : la sensibilité et la spécificité des scores globaux démontrent une capacité à identifier justement la présence et l'absence de troubles de l'alimentation et/ou de la déglutition chez le nouveau-né.

Hypothèse 2 : Les items de l'outil testé possèdent une bonne validité de critère : la sensibilité et la spécificité de chaque item démontrent une capacité à identifier justement la présence ou l'absence de troubles de l'alimentation et/ou de la déglutition chez le nouveau-né.

3 METHODOLOGIE

3.1 PARTICIPANTS

Afin d'étudier la validité de critère de la grille d'évaluation testée, nous nous sommes appuyé sur la population recrutée pour la phase III de l'étude OTBB3 menée par le Professeur M. Tauber. Il s'agissait de 11 nourrissons, 5 filles et 6 garçons, âgés de 1 à 4 mois au début de l'étude, recrutés entre mars 2020 et septembre 2021.

L'objet de l'étude était l'analyse en double aveugle de l'efficacité sur la succion d'un traitement par ocytocine. A chaque visite au Centre de Référence du Syndrome de Prader-Willi, une vidéo de tétée était réalisée, ainsi qu'une VRDG lors de la visite d'inclusion (V1), puis à un mois de traitement par ocytocine ou placebo (V4), et enfin à 6 mois du début de l'étude (V8).

Concernant la sélection des sujets pour l'étude OTBB3, ceux-ci devaient respecter les critères qui suivent (P. M. Tauber & Fichaux-Bourin, s. d.).

Critères d'inclusion :

- Nouveau-né ou nourrisson de sexe masculin ou féminin, présentant un SPW génétiquement confirmé
- Âgé de moins de 92 jours, avec une tolérance maximale de huit jours (pour les nourrissons prématurés, nés avant 37 semaines d'aménorrhée, un âge corrigé était appliqué)
- Consentement éclairé signé obtenu auprès des parents/titulaires de l'autorité parentale
- Parents souhaitant et acceptant de se conformer à toutes les procédures de l'étude

Critères d'exclusion :

- Nouveau-né ou nourrisson admis en service d'urgence pour des comorbidités engageant le pronostic vital (notamment anomalies respiratoires, cardiovasculaires ou neurologiques sévères)
- Nouveau-né ou nourrisson présentant un risque d'anomalie de conduction cardiaque (allongement de l'intervalle QT)
- Nouveau-né ou nourrisson non couvert par une assurance maladie
- Nouveau-né ou nourrisson présentant une hypersensibilité à l'ocytocine ou aux excipients du produit

- Nouveau-né ou nourrisson présentant une hypokaliémie cliniquement significative (à la discrétion du médecin)
- Nouveau-né ou nourrisson participant simultanément à une autre étude interventionnelle
- Nouveau-né ou nourrisson dont la situation des parents peut compromettre l'interprétation des résultats
- Nouveau-né ou nourrisson dont les parents refusent l'enregistrement vidéo, requis pour répondre à l'objectif principal de l'étude

Pour notre étude, les 11 patients de l'étude OTBB3 ont été retenus. Certaines vidéos de tétée ont cependant été éliminées. En effet, celles qui n'étaient pas doublées d'une VRDG du même jour de visite ne présentaient pas d'intérêt pour notre étude, de par l'impossibilité de comparer ces deux examens entre eux.

Pour les 10 premiers patients, 3 visites ont été gardées (V1, V4, V8). Le onzième patient n'ayant pu passer l'examen de VRDG à la 4ème visite, seules la première et la huitième ont été gardées (V1 et V8). Au total, 32 appariements vidéo tétée - VRDG seront analysés.

3.2 MATERIEL ET PROCEDURES

3.2.1 Procédure pour l'évaluation clinique

3.2.1.1 *Présentation de la grille d'observation clinique*

Comme cela a été présenté plus haut, la grille d'évaluation clinique étudiée se compose de cinq domaines s'intéressant aux différents temps d'alimentation du nourrisson.

Nous nous sommes appuyé sur la grille révisée par Emilie Leboucher pour l'étude de sa fidélité inter-juges, composée de 16 items, principalement centrés sur les volets déglutition et alimentation.

Cependant, notre étude reposant sur la comparaison de cette grille clinique à l'examen de la VRDG, nous avons réduit le nombre d'items étudiés à 8. Suite à l'appariement des items des deux outils en question (dont nous parlerons plus tard), nous n'avons effectivement gardé que ceux objectivables en VRDG et permettant d'évaluer les mécanismes physiopathologiques et leur sévérité.

Les 8 items ainsi sélectionnés sont présentés dans un tableau proposant 3 niveaux de sévérité : un premier niveau « normal » coté 1, un deuxième niveau « à risque » coté 2 et un troisième niveau « anormal » coté 3 (voir tableau 1 ci-dessous).

Ils concernent, d'une part, la phase orale de la déglutition, avec observation de la continence labiale, de la mandibule, du rythme de succion et de la langue.

D'autre part, la séquence de déglutition pharyngée est étudiée. Est alors observée la synchronisation déglutition-respiration.

Enfin, les trois derniers items permettent d'étudier la phase pharyngée et la protection des voies respiratoires en s'intéressant à la propulsion pharyngée, à la continence vélopharyngée et aux mécanismes d'expulsion (ANNEXE 2).

	NORMAL (1)	A RISQUE (2)	ANORMAL (3)	RESULTATS (1 / 2 / 3)
Item 1 : Continence labiale	Contenance labiale 1- Maintenu	2- Non maintenue	3- Absente	
Item 2 : Mouvements mandibulaires	Mouvements mandibulaires adaptés 1 - Mouvements adaptés	2 - Mouvements trop amples	3 - Mouvements inefficaces	
Item 3 : Coordination des mvts - Rythme de succion	Rythme de succion-déglutition pdt 2 min 1 - > 10 suctions-déglutitions	2 - < 10 suctions-déglutitions	3 - Pas de rythme	

Tableau 1 : Aperçu de la grille d'observation clinique utilisée pour les cotations de tétée

3.2.1.2 Protocole d'enregistrement vidéo

Dans cette étude, nous avons utilisé des vidéos de tétée provenant de l'étude OTBB3 du Pr M. Tauber traitant de l'effet de l'ocytocine sur les enfants porteurs du Syndrome de Prader-Willi (SPW).

Ces vidéos ont été enregistrées selon un protocole strict garantissant du mieux possible une collecte de données rigoureuse. Pour chaque enregistrement, un numéro d'anonymat était attribué.

Les nourrissons étaient filmés au moyen du même appareil numérique, par un personnel formé, dans les conditions les plus écologiques possibles. L'enfant était confortablement installé, le cou dégagé, nourri par l'un de ses parents avec son biberon habituel. Cet enregistrement se déroulait si possible dans un environnement calme, afin de pouvoir distinguer les bruits liés à la prise alimentaire des bruits environnants.

Si, pour les besoins de l'étude OTBB3, ces vidéos duraient entre 20 et 30 minutes, nous nous sommes contenté pour notre étude de l'observation des premières minutes uniquement. D'après la littérature scientifique, les deux premières minutes de tétée sont « les plus efficaces » (Longoni et al., 2018; Zarem et al., 2013). Nous avons étendu l'observation à 5 minutes afin d'avoir un aperçu plus global des capacités du nourrissons.

3.2.2 Procédure pour la vidéoradioscopie de la déglutition

3.2.2.1 *Présentation de la grille de vidéoradioscopie*

La grille d'évaluation utilisée pour coter les enregistrements vidéoradioscopiques des nourrissons a été élaborée par le Dr Fichaux-Bourin, et utilisée dans l'étude "The Use of Oxytocin to Improve Feeding and Social Skills in Infants With Prader-Willi Syndrom" (M. Tauber et al., 2017).

Elle se présente initialement sous la forme d'un tableau de 9 items, répartis en 3 domaines (phase orale, initiation et synchronisation de la déglutition pharyngée, phase pharyngée et protection des voies aériennes) et cotés de 1 à 5 points (avec 1 point : normal ; 2 points et plus : anormal, du moins au plus sévère).

Dans notre étude, nous n'avons gardé que 8 items, l'item "pauses respiratoires" n'étant pas appariale de manière pertinente avec ceux de la grille d'observation de tétée. C'est un item à caractère plutôt qualitatif, correspondant à l'endurance et difficilement objectivable en VRDG. Effectivement, cet examen propose des séquences d'observation très courtes du fait de son caractère irradiant. Les pauses sont donc observables sur un temps beaucoup plus restreint et moins représentatif que dans une situation de tétée. Elles n'ont pas la même importance et le même poids dans ces deux situations d'évaluation, nous n'avons donc pas jugé pertinent de conserver cet item.

Pour des raisons de simplification de calculs statistiques, les cotations ont été reprises et réparties de la manière suivante : le critère "normal" (initialement 1 point) était remplacé par la cotation 0 point et les critères "pathologique" (initialement 2 à 5 points) étaient remplacés par la cotation 1 point (voir tableau 2 ci-dessous). Cette manière binaire de coter, si elle perdait en précision concernant la sévérité de l'atteinte, nous permettait ainsi d'attribuer plus facilement au patient le statut "normal" ou "anormal". Cette attribution malade/sain est effectivement le rôle du Gold Standard (ici, la grille de VRDG), afin de permettre une comparaison entre celle-ci et l'outil testé (la grille d'observation clinique).

De cette manière, un score global supérieur ou égal à 1/8 était considéré comme anormal, tandis que seul le score 0/8 était reconnu normal.

Domaine	Composante	Observation	Résultats Normal (0 pt) / Anormal (1 pt)
Phase orale	Fermeture labiale	Continente	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)
	Mouvements mandibulaires	Stabilité	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)
	Mouvements linguaux	Nombre de mouvements de succion par déglutition	<input type="checkbox"/> 1 à 3 (0pt) <input type="checkbox"/> >3 ou 0 (1pt)
	Coordination des mouvements	Rythme de succion régulier	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)

Tableau 2 : Aperçu de la grille utilisée pour la cotation vidéoradioscopique

Les 3 mêmes domaines ont été gardés, évaluent chacun un temps différent de la séquence de déglutition :

- La phase orale : il s'agit d'observer la qualité de la fermeture labiale, des mouvements mandibulaires, des mouvements linguaux et du rythme de succion de l'enfant.
- L'initiation et la synchronisation de la déglutition pharyngée : on s'intéresse à la localisation du bolus lors de ce temps.
- La phase pharyngée et la protection des voies aériennes : sont étudiées la qualité de la propulsion pharyngée et de la protection des voies aériennes supérieures (continence vélopharyngée et pénétrations et/ou inhalations) (ANNEXE 3).

3.2.2.2 Protocole d'enregistrement vidéoradioscopique

De la même manière qu'avec les enregistrements vidéo, nous nous sommes appuyé sur les enregistrements de VRDG provenant de l'étude OTBB3.

Le protocole pour la réalisation de ces enregistrements était détaillé dans un document transmis aux professionnels concernés et formés à la passation de cet examen. Celui-ci était toujours réalisé en présence du médecin spécialisé en déglutition.

Il comprenait des informations précises sur :

- L'installation du nourrisson : bien calé dans un matelas coquille légèrement basculé afin de reproduire le plus fidèlement possible les bras du parent.
- L'administration du bolus baryté : proportion du produit de contraste, biberon et tétine utilisés ainsi que leur positionnement.

- L'irradiation : instructions concernant le champ de prise et la durée afin de limiter la dose qui figurait pour chaque examen, exprimée en mGy ou μGy^2 .

Les données recueillies étaient anonymisées et transférées sur une plateforme de lecture sécurisée (Centre Référent SPW, 2020).

3.2.3 Appariement des items des deux grilles

Les items de chacune des deux grilles ont donc été étudiés et mis en concordance dans un tableau de 8 items, numérotés de 1 à 8. Par souci de cohérence, nous avons décidé de les présenter par domaine. Chacun des 3 domaines ainsi définis concernait un temps de la déglutition : la phase orale, l'initiation et la synchronisation de la déglutition pharyngée, la phase pharyngée et la protection des voies aériennes (ANNEXE 4).

3.2.3.1 *Item 1 : Continence labiale*

Le premier item de la grille d'observation clinique cherche à déterminer si la continence des lèvres est bien maintenue durant la tétée. C'est également ce que recherche l'item « fermeture labiale » de la grille vidéoradioscopique.

Nous avons donc apparié ces deux items ensemble sous la dénomination : « Item 1 : Continence labiale ».

3.2.3.2 *Item 2 : Mouvements mandibulaires*

Un des items de la grille d'observation clinique questionne l'adaptation des mouvements mandibulaires à la succion, tandis que dans la grille de VRDG, un autre traite de leur stabilité.

Or, la stabilité des mouvements mandibulaires témoignant d'une succion efficace (Arvedson et al., 2019), nous avons mis ces deux items en concordance sous l'appellation « Item 2 : Mouvements mandibulaires ».

3.2.3.3 *Item 3 : Rythme de succion*

Un item de chaque grille traite de la coordination des mouvements par l'observation de la capacité du nourrisson à tenir un rythme de succion-déglutition de manière régulière.

Ils ont donc été appariés pour former l'item 3 : « Rythme de succion ».

3.2.3.4 *Item 4 : Mouvements linguaux*

L'item « mouvements linguaux » de la grille de VRDG permet de relever le nombre de mouvements de succion par déglutition. 1 à 3 mouvements linguaux sont jugés dans la norme, alors que dans le cas d'une absence ou de la réalisation de plus de 3 mouvements, on cote l'item en « anormal ». Dans la grille d'observation clinique, un item évalue la synchronisation succion-déglutition en relevant de la même manière le nombre de mouvements de succion par déglutition.

Ces deux items ont donc été appariés sous l'appellation « Item 4 : Mouvements linguaux ».

3.2.3.5 *Item 5 : Initiation de la déglutition pharyngée*

Dans la grille de VRDG, l'item « localisation du bolus » au moment de l'initiation de la déglutition pharyngée témoigne de la qualité et du bon déroulé de cette dernière.

Du côté de la grille d'observation, l'item intitulé « synchronisation déglutition-respiration » s'intéresse en réalité à la fluidité de la séquence et ainsi à la bonne initiation de la déglutition pharyngée. Une désynchronisation déglutition-respiration serait la conséquence d'un retard de déclenchement.

Ces deux items traitent donc de l'analyse du même mécanisme, à savoir : « l'initiation de la déglutition pharyngée ». Nous les avons donc appariés sous cette dénomination, pour l'item 5.

3.2.3.6 *Item 6 : Propulsion pharyngée*

Lors du temps pharyngé, le type de bruits (respiratoires, pharyngés, encombrements) ainsi que la présence ou l'absence de stases permettent déterminer la qualité de la propulsion pharyngée.

L'item "bruits respiratoires" de la grille d'observation clinique et l'item "présence de stases" de la grille de VRDG ont donc été appariés sous l'appellation "Item 6 : Propulsion pharyngée" (Arvedson et al., 2019).

3.2.3.7 *Item 7 : Continence vélopharyngée*

L'un des items de la grille d'observation clinique se préoccupe de la présence ou non d'écoulement nasal de lait. Du côté de la grille vidéoradioscopique, un item concerne l'existence éventuelle d'un reflux nasopharyngé.

Il s'agit, dans ces deux items, de l'observation d'une seule et même caractéristique de la déglutition. Ils ont donc été appariés, formant l'item 7 : « Continence vélopharyngée ».

3.2.3.8 *Item 8 : Protection des voies respiratoires*

Les derniers items de chacune des deux grilles d'observation clinique et de VRDG concernent respectivement les mécanismes d'expulsion et les potentielles pénétrations et/ou inhalations.

Dans ces deux situations, la question se pose de la qualité de la protection des voies respiratoires. Les deux items ont ainsi été regroupés pour former l'item 8 : "Protection des voies respiratoires".

3.2.3.9 *Appariement des systèmes de notation*

Comme cela a été décrit plus haut, les deux grilles ne présentent pas le même système de notation. La grille d'observation clinique comprend trois critères (normal, à risque, anormal), tandis que la grille de VRDG est binaire (normal, anormal).

La question s'est donc posée de l'intégration du critère « à risque » de la grille testée au critère « normal » ou bien au critère « anormal ». Nous nous sommes demandé si l'on considérait que l'obtention de ce critère à un item impliquait une simple surveillance (associé au critère « normal ») ou bien nécessitait la réalisation d'un examen de VRDG (associé au critère « anormal »).

C'est cette question que nous nous sommes posée en testant ces deux cas de figure. Deux protocoles ont ainsi été créés afin de déterminer lequel était le plus pertinent et le mieux corrélé à la grille de VRDG.

Dans le protocole A, le critère "à risque" est traité avec le critère "anormal". Dans le protocole B, il est traité avec le critère "normal" (voir tableau 3 ci-dessous).

		Grille Observation Clinique				Grille VRDG			
		ANALYSE SCORES GLOBAUX		ANALYSE PAR ITEM		ANALYSE SCORES GLOBAUX		ANALYSE PAR ITEM	
DOMAINE		NORMAL (0 pt ou-) / ANORMAL (1 pt ou +)	RESULTAT 0 pt (1) / 1 pt (2-3) <i>Protocole A</i>	RESULTAT 1-2 (0 pt) / 3 (1 pt) <i>Protocole B</i>	RESULTAT T- (1) / Tr- (-3) <i>Protocole 2</i>	RESULTAT T- (-1) / Tr- (-3) <i>Protocole 3</i>	NORMAL (0 pt ou-) / ANORMAL (1 pt ou +)	RESULTAT 0 pt / 1 pt	RESULTAT M- / M+
PHASE ORALE	Item 1 : Continence labiale	Continence labiale 1- Maintenu (0) 2- Non maintenu (0/1) 3- Absence (1)					Continence labiale Oui (0 pt) Non (1 pt)		
	Item 2 : Mvts mandibulaires	Mvts mandibulaires adaptés 1- Mouvements adaptés (0) 2- Mouvements trop amples (0/1) 3- Mouvements inefficaces (1)					Mvts stables Oui (0 pt) Non (1 pt)		
	Item 3 : Rythme de succion	Rythme de succion-déglutition pdt 2 min 1- > 10 suctions-déglutitions (0) 2- < 10 suctions-déglutitions (0/1) 3- Pas de rythme (1)					Rythme de succion régulier Oui (0 pt) Non (1 pt)		

Tableau 3 : Aperçu du tableau d'appariement pour l'analyse globale et par item - Protocoles A et B

3.2.4 Recueil et préparation des données

Pour des raisons de rigueur statistique, les cotations des deux grilles devaient être effectuées par le même examinateur. Or, comme expliqué plus haut, seul un professionnel qualifié était en mesure de coter les VRDG. C'est pour cette raison que toutes les cotations ont été effectuées par le Dr Pascale Fichaux-Bourin.

Afin de respecter au mieux les conditions de cotation en direct dans lesquelles cet outil sera utilisé, les vidéos auront été visionnées au maximum deux fois par l'examinateur qui les cote.

Ces données étaient recueillies dans des fichiers Excel, puis recoupées au sein d'un même fichier, dans 32 tableaux d'appariement correspondant aux 32 duos de vidéos traitées (ANNEXE 4).

3.3 ANALYSE DES DONNEES - STATISTIQUES

3.3.1 Démarche générale

Pour répondre à notre problématique, poser nos hypothèses et construire notre méthodologie, nous nous sommes appuyé sur les normes psychométriques COSMIN, citées plus haut.

Les vidéos de tétée et l'examen de VRDG étant effectués, pour chaque enfant, sur la même journée de visite, nous nous sommes intéressé, pour l'étude de la validité de critère, à la validité concurrente.

Afin d'étudier de manière plus complète la validité de critère de notre grille d'observation, nous avons décidé de l'analyser sous deux angles de vue différents.

Dans un premier temps, nous l'étudierons avec un regard global, en nous intéressant à la comparaison des scores totaux de la grille testée avec ceux du gold standard.

Dans un second temps, nous nous intéresserons plus précisément à la validité de chacun des items de la grille testée.

3.3.2 Analyse des scores globaux

Les deux grilles se composent de 8 items, chacun coté 0 (normal) ou 1 (anormal). Nous avons donc deux scores notés /8 pour chacun des 32 appariements. Afin de comparer rigoureusement ces deux scores globaux, les valeurs seuils de ceux-ci devait être clairement identifiées.

Celle-ci était connue pour la grille de VRDG qui n'acceptait aucun item positif et classait donc tous les scores de 1/8 ou plus dans la catégorie « anormal ».

Concernant la grille d'observation clinique, qui est un test quantitatif, l'enjeu était justement de déterminer le seuil optimal de ce test pour le considérer comme normal (T-) ou anormal (T+).

C'est pourquoi nous avons construit, pour chacun des deux protocoles, les courbes ROC (Receiver Operating Characteristic), permettant cette décision du seuil et la comparaison des courbes.

A partir de ce seuil ainsi fixé, plusieurs paramètres, que nous détaillerons plus tard, étaient observables : l'air sous la courbe (en anglais AUC : « Area Under Curb »), les qualités intrinsèques avec l'étude la Sensibilité (Se) et de la Spécificité (Spe) et les qualités extrinsèques du test avec le calcul des Valeurs Prédicatives Positive et Négative (VPP et VPN).

3.3.3 Analyse des données item par item

L'analyse des données de chaque item est une analyse dichotomique. Pour chacun de ces 8 items, un critère « normal » ou « anormal » était attribué par la grille d'observation testée (T- ou T+) ainsi que par l'examen de VRDG (M- ou M+).

A partir de ces données, nous avons donc construit pour chacun des items un tableau de contingence à double entrée (voir tableau 4 ci-dessous). A partir de ce tableau, sensibilité et spécificité ont été observés. L'indice de Youden a également été calculé, afin d'étudier la relation entre ces deux premiers paramètres, et déterminer ainsi la validité de critère de chacun de ces items.

Les VPP et VPN, habituellement calculées dans cette situation, ont été laissées de côté pour cette analyse par item, de par l'impossibilité d'en respecter les conditions d'application. Ces valeurs dépendent effectivement de la prévalence de la maladie ou du trouble étudié, dans la population générale. Il était impossible, dans notre cas, de déterminer la prévalence générale des difficultés spécifiques ciblées par chaque item.

		VRDG		
		M+	M-	Total
Grille	T+	VP	FP	VP+FP
	T-	FN	VN	FN+VN
	Total	VP+FN	FP+VN	TOTAL

Légende

- **M+** : résultat anormal selon la VRDG
- **M-** : résultat normal selon la VRDG
- **T+** : résultat anormal selon la grille d'observation testée
- **T-** : résultat normal selon la grille d'observation testée
- **VP** (Vrais Positifs) : personnes ayant obtenu un résultat positif avec les deux instruments de mesure
- **FP** (Faux Positifs) : personnes ayant obtenu un résultat positif avec la grille testée et un résultat négatif avec la VRDG de référence
- **FN** (Faux Négatifs) : personnes ayant obtenu un résultat négatif avec la grille testée et un résultat positif avec la VRDG de référence
- **VN** (Vrais Négatifs) : personnes ayant obtenu un résultat négatif avec les deux instruments de mesure

Tableau 4 : Tableau de contingence entre la grille d'observation et la grille de VRDG

3.3.4 Les différents paramètres étudiés

3.3.4.1 L'aire sous la courbe (AUC)

L'aire sous la courbe permet d'estimer le pouvoir discriminant global du test. Sa valeur est comprise entre 0,5 et 1. Un test est considéré comme peu discriminant pour $0,5 \leq AUC < 0,7$, moyennement discriminant pour $0,7 \leq AUC < 0,9$, très discriminant pour $0,9 \leq AUC < 1$, et parfait pour $AUC = 1$.

En pratique, une AUC supérieure à 0,7 sera considérée comme valable pour un test diagnostic (Delacour et al., 2005).

3.3.4.2 Les qualités intrinsèques

Les qualités intrinsèques d'un test sont définies et calculées en conditions expérimentales et sont donc indépendantes du type de personnes testées. Elles comprennent la sensibilité, la spécificité et l'indice de Youden (Delacour et al., 2005; Tétreault & Guillez, 2014).

La Sensibilité (Se) représente la probabilité de présenter un test positif (T+) lorsque le patient est malade (M+). Elle permet donc d'identifier correctement la présence d'une maladie ou d'un trouble chez un individu.

La Spécificité (Spe) représente pour sa part la probabilité de présenter un test négatif (T-) lorsque le patient est sain (M-). Elle permet donc d'identifier correctement l'absence d'une maladie ou d'un trouble chez un individu.

Ces paramètres sont compris entre 0 et 1 (ou parfois exprimés en pourcentage). Plus ils seront proches de 1 (ou 100%), plus l'item étudié sera sensible/spécifique. Une sensibilité à 0,5 (ou 50%) signifierait qu'il y a autant de vrais positifs (VP) que de faux négatifs (FN), et donc que l'item néglige un individu malade sur deux. De même, une spécificité à 0,5 (ou 50%) signifierait qu'il y a autant de vrais négatifs (VN) que de faux positifs (FP), et donc que l'item néglige un individu sain sur deux. Un item parfait aurait ainsi une sensibilité et une spécificité toutes deux à 1.

Cependant, en pratique, un gain de sensibilité est obtenu contre une perte de spécificité, et inversement. L'interprétation doit donc se faire de manière clinique, et, selon les résultats ainsi obtenus, les objectifs futurs du test peuvent être déterminés. Dans le cas d'une meilleure sensibilité, le test aura plutôt une valeur de dépistage. Dans le cas d'une meilleure spécificité, il sera plutôt à valeur diagnostique.

Concernant l'analyse par item, nous avons donc décidé de nous baser sur la valeur de l'indice de Youden pour établir le caractère discriminant ou non discriminant de chaque item.

L'indice de Youden (J) est une autre mesure permettant de déterminer la précision de la méthode de diagnostic. Cet indicateur de synthèse présente l'avantage de mettre en lien ces deux paramètres de sensibilité et spécificité, et complète de cette manière notre regard sur la validité de critère de chacun de ces items.

Il est compris entre -1 et 1. Pour $J < 0$, le test est jugé inefficace. Pour $J = 0$, il néglige un individu sain sur deux, et un individu malade sur deux, témoignant également d'une efficacité insuffisante. Pour $J = 1$, il identifie justement tous les patients malades et sains comme tels.

En l'absence de consensus quant au seuil d'acceptabilité de ce paramètre, nous avons décidé arbitrairement de le fixer à 0,5, pour une meilleure fiabilité de nos résultats. Ainsi, les items présentant un indice de Youden $J \geq 0,5$ étaient considérés comme témoignant d'une bonne validité de critère.

3.3.4.3 Les qualités extrinsèques

Les qualités extrinsèques d'un test sont relatives à l'utilisation du test pour une population donnée et diffèrent selon les caractéristiques de la population testée. Elles sont définies et calculées en situation de dépistage et permettent d'apprécier la pertinence de l'utilisation du test dans cette population précise. Elles comprennent la Valeur Prédictive Positive (VPP) et la Valeur Prédictive Négative (VPN).

La Valeur Prédictive Positive (VPP) est la probabilité que l'individu soit réellement malade si le test est positif. Elle représente la confiance accordée à un test positif.

La Valeur Prédictive Négative (VPN) est la probabilité que l'individu soit réellement sain si le test est négatif. Elle représente la confiance accordée à un test négatif.

Toutes deux sont généralement comprises entre 0 et 1 (ou exprimées en pourcentage). Comme pour la sensibilité et la spécificité, plus on approche de 1 (ou 100%), meilleure est la capacité diagnostique de l'outil testé.

Cependant, comme évoqué plus haut, ces valeurs ont pour particularité de dépendre de la prévalence du trouble ou de la maladie étudiée dans la population générale. Ainsi, il est nécessaire pour ces calculs de connaître cette prévalence générale afin de la comparer à celle retrouvée dans notre échantillon. Si l'on retrouve une différence entre ces deux grandeurs, ce rapport sera à prendre en compte dans le calcul de ces paramètres. Si la répartition d'individus malades est la même entre l'échantillon et la population générale, le calcul s'effectuera plus simplement, avec les valeurs provenant du tableau de contingence.

4 RESULTATS

4.1 ANALYSE DES SCORES GLOBAUX

Les analyses ont été effectuées avec le logiciel Excel et le logiciel XLSTAT.

4.1.1 Résultats du protocole A

4.1.1.1 Statistiques descriptives

Pour rappel, le protocole A proposait de comptabiliser le critère médian « à risque » de la grille d'observation testée en critère « anormal ». Nous avons ainsi coté 0 point pour le critère « normal » contre 1 point pour les critères « à risque » et « anormal ».

Sur les 32 vidéos de VRDG, seules 4 ont obtenu un score « normal » (M-) (score de 0/8), contre 28 vidéos cotées « anormal » (M+). Nous avons donc, dans notre échantillon, une prévalence de 88%.

Concernant la cotation des tétées au moyen de la grille d'observation clinique, les scores étaient répartis de 0/8 à 7/8 (voir figure 1 ci-dessous), avec une moyenne de 2,91/8 et un écart-type de 2,04 (ANNEXE 5).

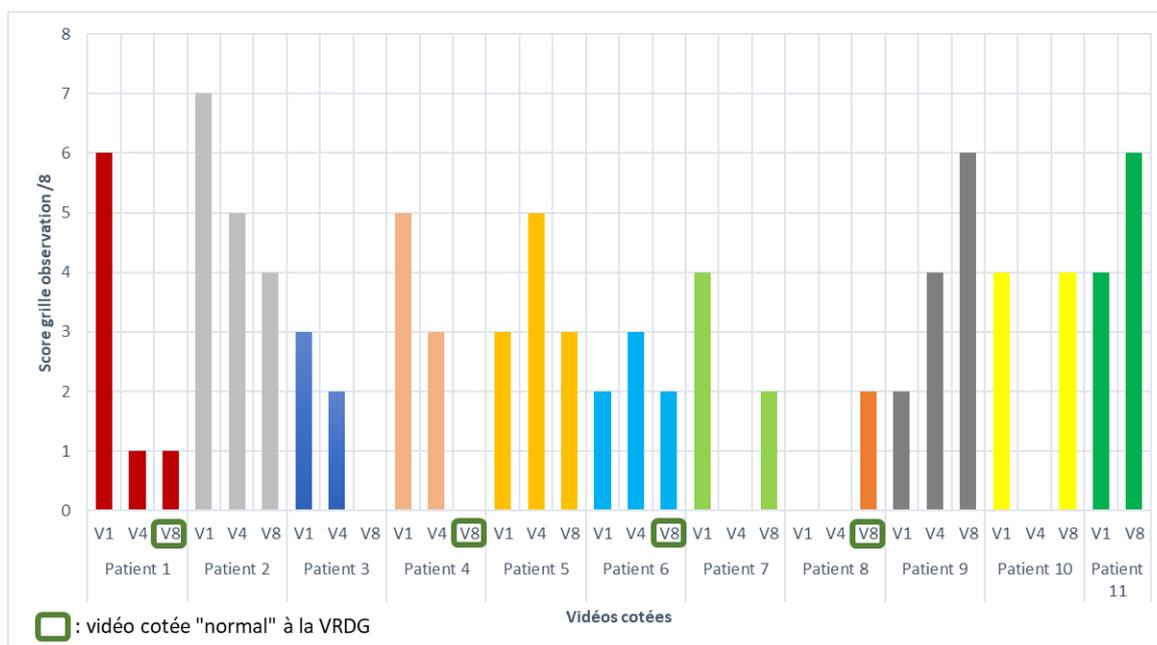


Figure 1 : Notation de chaque vidéo via la grille d'observation clinique - Protocole A

4.1.1.2 Courbe ROC

4.1.1.2.1 Détermination de la valeur seuil optimale

Selon Delacour et al. (2005) : « la courbe ROC est une représentation graphique de la relation existant entre la sensibilité et la spécificité d'un test, calculée pour toutes les valeurs seuils possibles. Elle permet la détermination et la comparaison des performances diagnostiques de plusieurs tests à l'aide de l'évaluation des aires sous la courbe. ». Dans le cas de notre étude, l'analyse nous permettra ainsi une comparaison entre les protocoles A et B de la grille d'observation clinique testée.

Ainsi, l'analyse ROC permet également la détermination de la valeur seuil optimale dans un test quantitatif comme le nôtre. Après construction de la courbe grâce à la proportion de faux positifs et de vrais positifs (voir figure 2 ci-dessous), cette valeur seuil est obtenue en recherchant le point de la courbe le plus éloigné de la diagonale qui, elle, représente un test d'apport nul.

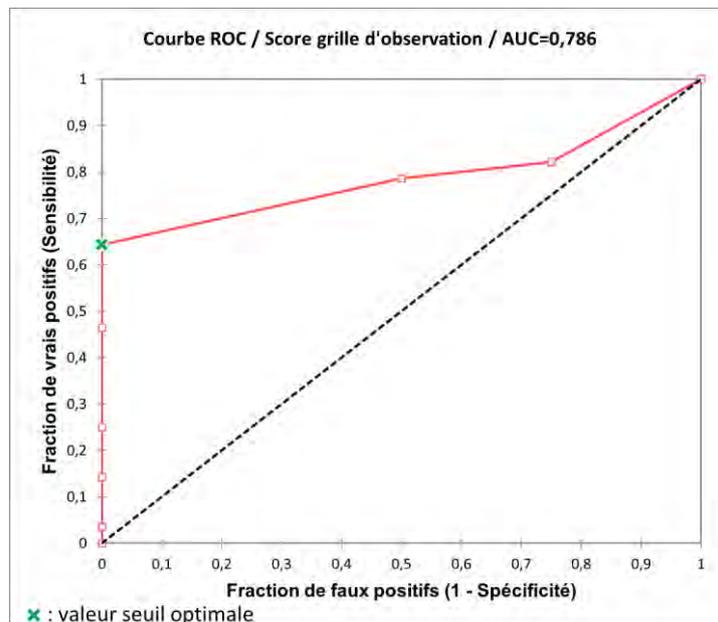


Figure 2 : Courbe ROC obtenue pour les scores de la grille d'observation clinique - Protocole A

Grâce au logiciel XLSTAT, nous avons pu déterminer la valeur seuil optimale, correspondant à ce point de la courbe, à attribuer à la grille. Ainsi, un score inférieur à 3/8 était considéré comme « normal » (T-) tandis qu'un score supérieur ou égal à 3/8 était considéré comme « anormal » (T+) (ANNEXE 6).

4.1.1.2.2 Tableau de contingence et paramètres associés

A partir de cette valeur seuil optimale de 3/8, nous avons pu construire un tableau de contingence permettant de représenter la proportion de Vrais Positifs (VP), Faux Positifs (FP), Vrais Négatifs (VN) et Faux Négatifs (FN) (voir tableau 5 ci-dessous).

		VRDG		
		M+	M-	Total
Grille	T+	18	0	18
	T-	10	4	14
	Total	28	4	32

Légende

- **M+** : résultat anormal selon la VRDG
- **M-** : résultat normal selon la VRDG
- **T+** : résultat anormal selon la grille d'observation testée
- **T-** : résultat normal selon la grille d'observation testée

Tableau 5 : Tableau de contingence pour la valeur seuil de 3/8 - Protocole A

Une fois ces valeurs établies, les paramètres correspondant aux qualités intrinsèques et extrinsèques du test ont pu être calculés.

La sensibilité était de 0,64 avec un intervalle de confiance à 95% de [0,46 ; 0,79]. Il y a donc 64% de chances de présenter un test positif lorsque le patient est malade. La spécificité était de 1,00 avec un intervalle de confiance à 95% de [0,45 ; 1,00]. Il y a donc 100% de chances de présenter un test négatif lorsque le patient est sain. L'indice de Youden, quant à lui, était de 0,64. Ces paramètres sont tous supérieurs à 0,5 ce qui signifie que, sans repérer fidèlement tous les malades et tous les sains, ce test néglige moins d'un individu malade/sain sur deux.

Les troubles de la déglutition sont des difficultés quasi-systématiques dans la population générale d'enfants porteurs du Syndrome de Prader-Willi (*Protocole National de Diagnostic et de Soins (PNDS) - Syndrome Prader-Willi, 2021*). Dans notre échantillon, 88% des enfants sont porteurs de troubles de la déglutition. Nous avons donc fait le choix de considérer que ces deux prévalences étaient suffisamment proches pour pouvoir calculer la VPP et la VPN.

La VPP était de 1,00, autrement dit, il y a 100% de chances que l'enfant ait réellement des troubles de la déglutition lorsque le test est positif. La VPN était de 0,29 avec un intervalle de confiance à 95% de [0,05 ; 0,52]. Autrement dit, il y a 29% de chances que l'enfant n'ait aucun trouble dysphagique lorsque le test est négatif.

La construction de la courbe ROC a également permis le calcul de l'AUC. Celle-ci était de 0,79 (> 0,7), ce qui témoigne ainsi d'un test moyennement discriminant selon le protocole A.

4.1.2 Résultats du protocole B

4.1.2.1 Statistiques descriptives

Pour rappel, le protocole B proposait quant à lui de comptabiliser le critère médian « à risque » de la grille d'observation testée en critère « normal ». Nous avons ainsi coté 0 point pour les critères « normal » et « à risque » contre 1 point pour le critère « anormal ».

Ce second protocole ne jouant pas sur les résultats de la VRDG, les 4 mêmes vidéos étaient jugées « normales » (M-), et la prévalence restait de 88% dans notre échantillon.

Concernant la cotation des tétées au moyen de la grille d'observation clinique, les scores étaient répartis de 0/8 à 3/8 (voir figure 3 ci-dessous), avec une moyenne de 0,75/8 et un écart-type de 1,02 (ANNEXE 5).

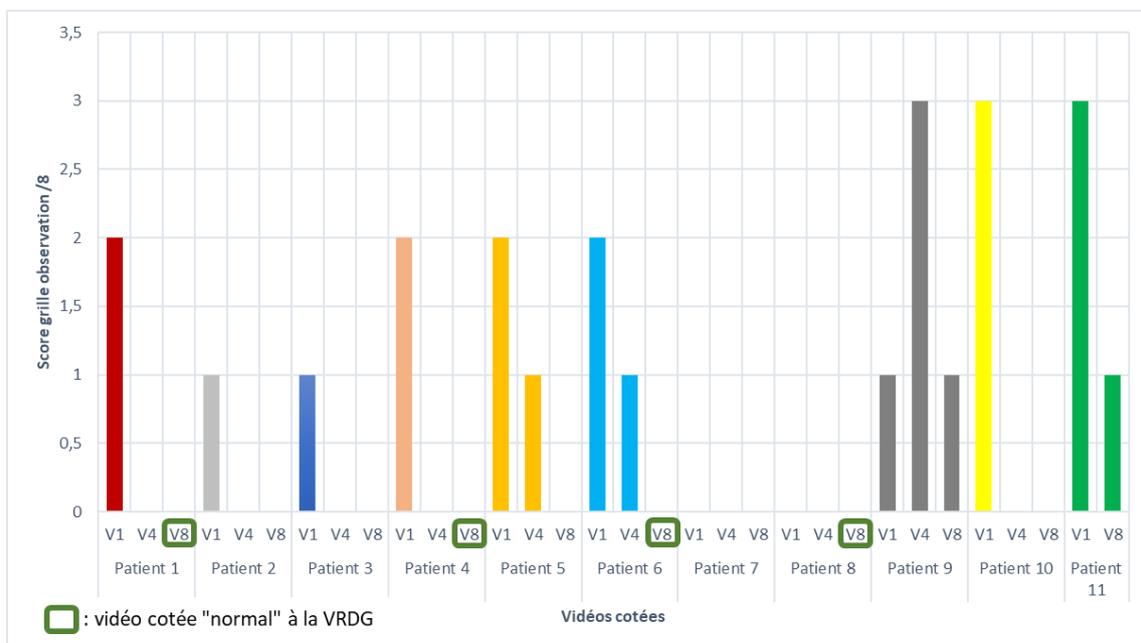


Figure 3 : Notation de chaque vidéo via la grille d'observation clinique - Protocole B

4.1.2.2 Courbes ROC

4.1.2.2.1 Détermination de la valeur seuil optimale

De la même manière que pour le protocole A, le logiciel XLSTAT nous a permis de tracer la courbe ROC (voir figure 4 ci-dessous) et de déterminer, pour le protocole B, la valeur seuil optimale à attribuer à la grille, correspondant à ce point de la courbe le plus éloigné de la diagonale. Ainsi, pour ce protocole B, un score inférieur à 1/8 était considéré comme « normal » (T-) tandis qu'un score supérieur ou égal à 1/8 était considéré comme « anormal » (T+) (ANNEXE 6).

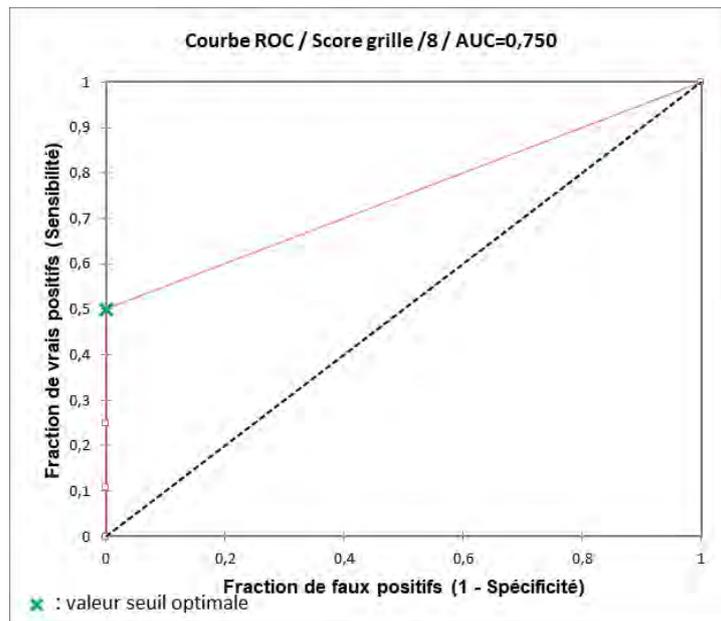


Figure 4 : Courbe ROC obtenue pour les scores de la grille d'observation clinique - Protocole B

4.1.2.2.2 Tableau de contingence et paramètres associés

A partir de cette valeur seuil optimale de 1/8, nous avons pu construire un tableau de contingence permettant de représenter la proportion de Vrais Positifs (VP), Faux Positifs (FP), Vrais Négatifs (VN) et Faux Négatifs (FN) (voir tableau 6 ci-dessous).

		VRDG		
		M+	M-	Total
Grille	T+	14	0	14
	T-	14	4	18
	Total	28	4	32

Légende

- **M+** : résultat anormal selon la VRDG
- **M-** : résultat normal selon la VRDG
- **T+** : résultat anormal selon la grille d'observation testée
- **T-** : résultat normal selon la grille d'observation testée

Tableau 6 : Tableau de contingence pour la valeur seuil de 1/8 - Protocole B

Une fois ces valeurs établies, les paramètres correspondant aux qualités intrinsèques et extrinsèques du test ont pu être calculés.

La sensibilité était de 0,50 avec un intervalle de confiance à 95% de [0,33 ; 0,67]. Il y a donc 50% de chances de présenter un test positif lorsque le patient est malade. La spécificité était de 1,00 avec un intervalle de confiance à 95% de [0,45 ; 1,00]. Il y a donc 100% de chances de présenter un test négatif lorsque le patient est sain. L'indice de Youden, quant à lui, était de 0,50. Si la spécificité présente un bon pourcentage, ce n'est pas le cas de la sensibilité et ainsi de la relation entre les deux que représente l'indice de Youden. Une sensibilité de 0,5 signifie en effet que ce test néglige un individu malade sur deux.

La VPP était de 1,00, autrement dit, il y a 100% de chances que l'enfant ait réellement des troubles de la déglutition lorsque le test est positif. La VPN était de 0,22 avec un intervalle de confiance à 95% de [0,03 ; 0,41]. Autrement dit, il y a 22% de chances que l'enfant n'ait aucun trouble dysphagique lorsque le test est négatif.

La construction de la courbe ROC a également permis le calcul de l'AUC. Celle-ci était de 0,75 (> 0,7), ce qui témoigne ainsi d'un test moyennement discriminant selon le protocole B.

4.1.3 Comparaison des deux protocoles

L'objectif de la décomposition en deux protocoles de la cotation de la grille d'observation clinique était d'identifier le plus pertinent d'entre eux. A posteriori, cela devait ainsi permettre de déterminer le mode de cotation à utiliser à l'avenir pour cet outil.

Nous avons donc jugé pertinent de comparer différents paramètres de ces deux protocoles.

Le graphique ci-dessous (voir figure 5) nous montre que, pour un même taux de VN et de FP, le protocole A présente une proportion plus élevée de VP (et ainsi plus faible de FN) que le protocole B. Du point de vue de ces paramètres, le protocole A semble donc plus intéressant.

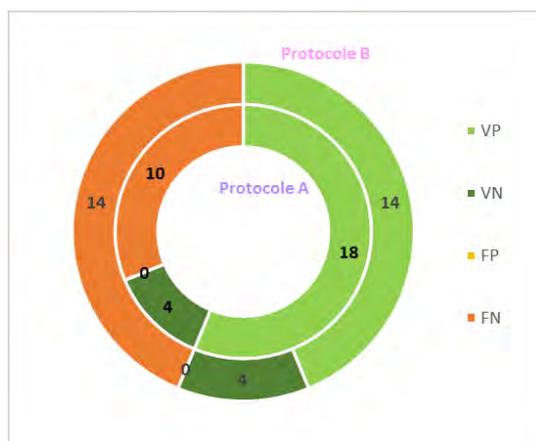


Figure 5 : Comparaison des répartitions du nombre de Vrais Positifs, Vrais Négatifs, Faux Positifs et Faux Négatifs pour chaque protocole

Le graphique ci-dessous (voir figure 6) montre que, pour une spécificité équivalente, le protocole A présente une meilleure sensibilité. Concernant ces paramètres, le protocole A présente ici aussi des qualités plus avantageuses.

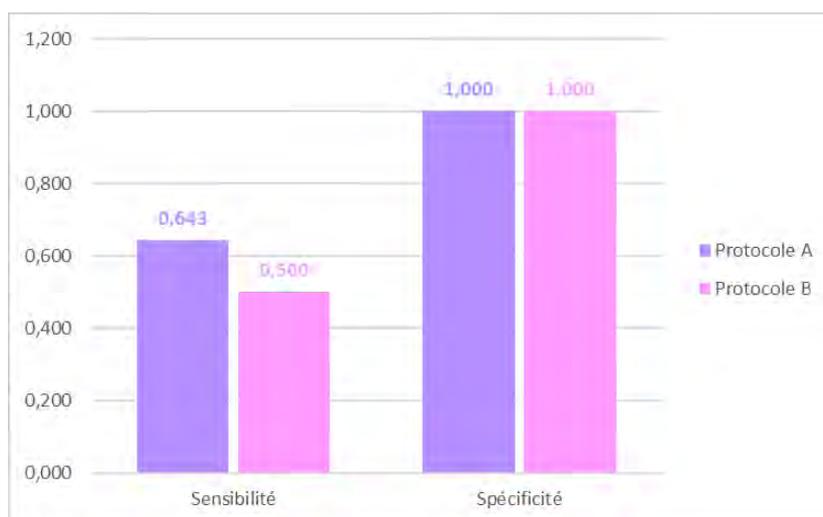


Figure 6 : Comparaison des valeurs de la Sensibilité et de la Spécificité des deux protocoles

En outre, l'indice de Youden dépend des valeurs de ces derniers paramètres. Il présente ainsi une valeur plus élevée pour le protocole A que pour le B ($J(A) = 0,64$ contre $J(B) = 0,50$).

De plus, ce graphique (voir figure 7 ci-dessous) nous indique que, pour une même VPP, le protocole A présente une fois encore une valeur plus élevée pour le deuxième paramètre, ici la VPN.

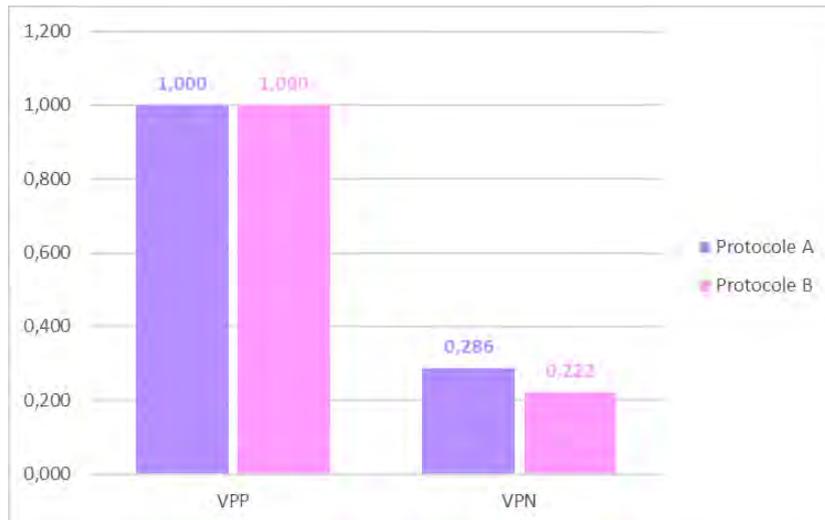


Figure 7 : Comparaison des valeurs de la VPP et de la VPN entre les deux protocoles

Enfin, la comparaison des aires sous la courbe (AUC) pour les deux protocoles nous permet de constater une nouvelle fois les meilleures capacités diagnostiques du protocole A par rapport au protocole B (AUC(A) = 0,79 contre AUC(B) = 0,75).

4.2 ANALYSE DES RESULTATS ITEM PAR ITEM

4.2.1 Résultats du protocole A

Pour chaque item, un tableau de contingence a été tracé, permettant de relever le nombre de VP, VN, FP et FN par item, pour le protocole A (ANNEXE 7). Par souci de clarté, ces valeurs ont été représentées au moyen d'un graphique à barres (voir figure 8 ci-dessous).

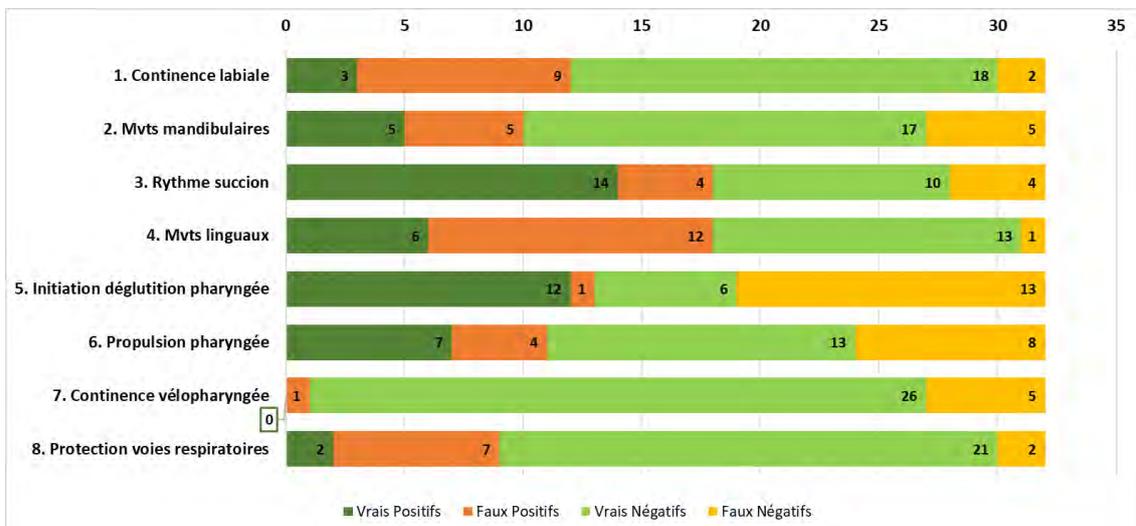


Figure 8 : Répartition par item du nombre de VP, VN, FP et FN - Protocole A

A partir de ces données, la sensibilité, la spécificité et l'indice de Youden de chacun de ces items ont pu être calculés (ANNEXE 7). Nous les avons regroupés dans le tableau qui suit (tableau 7).

Item n°	Protocole A				
	Se	IC(Se)95%	Spe	IC(Spe)95%	J
1. Continence labiale	0,6	[0,17 ; 1,02]	0,67	[0,49 ; 0,84]	0,27
2. Mvts mandibulaires	0,5	[0,19 ; 0,81]	0,77	[0,60 ; 0,95]	0,27
3. Rythme succion	0,78	[0,59 ; 0,97]	0,71	[0,48 ; 0,95]	0,49
4. Mvts linguaux	0,86	[0,60 ; 1,12]	0,52	[0,32 ; 0,72]	0,38
5. Initiation déglutition pharyngée	0,48	[0,28 ; 0,68]	0,86	[0,60 ; 1,12]	0,34
6. Propulsion pharyngée	0,47	[0,21 ; 0,72]	0,76	[0,56 ; 0,97]	0,23
7. Continence vélopharyngée	0	N.C.	0,96	[0,89 ; 1,03]	-0,04
8. Protection voies respiratoires	0,5	[0,01 ; 0,99]	0,75	[0,59 ; 0,91]	0,25

N.C. : Non Calculable

Tableau 7 : Sensibilité (Se), Spécificité (Spe) et indice de Youden (J) de chaque item - Protocole A

Selon l'indice de Youden, les items les plus fiables sont donc, en première place, l'item 3 « Rythme de succion-déglutition » (J = 0,49), puis l'item 4 « Mouvements linguaux » (J = 0,38) et l'item 5 « Initiation de la déglutition pharyngée » (J = 0,34). Se situent ensuite à égalité les items 1 « Continence labiale » (J = 0,27) et 2 « Mouvements mandibulaires » (J = 0,27), puis l'item 8 « Protection des voies respiratoires » (J = 0,25) suivi de l'item 6 « Propulsion pharyngée » (J = 0,23). Concernant l'item 7 « Continence vélopharyngée », les résultats ne sont pas concluants (J = -0,04). L'indice de Youden obtenu est effectivement négatif, ce qui témoigne de l'inefficacité diagnostique de cet item.

Cependant, selon notre seuil placé à 0,5, aucun de ces items n'atteste d'une bonne validité de critère.

De plus, pour ce protocole A, la spécificité est supérieure à la sensibilité pour 6 items sur les 8 (voir figure 9 ci-dessous).

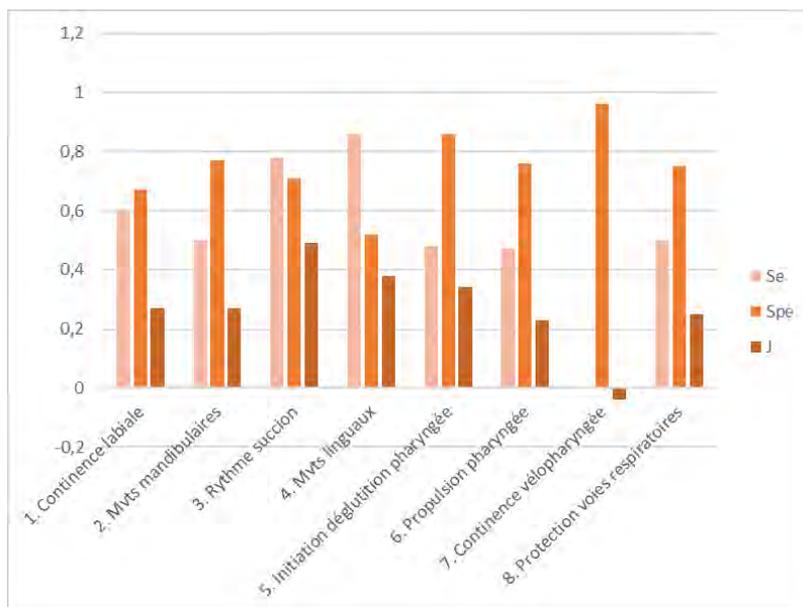


Figure 9 : Représentation de la Sensibilité (Se), de la Spécificité (Spe) et de l'indice de Youden (J) pour chaque item – Protocole A

4.2.2 Résultats du protocole B

Pour chacun des 8 items, dans le cadre du protocole B, un tableau de contingence a été tracé, permettant de relever le nombre de VP, VN, FP et FN par item (ANNEXE 7). Par souci de clarté, ces valeurs ont été représentées au moyen d'un graphique à barres (voir figure 10 ci-dessous).



Figure 10 : Répartition par item du nombre de VP, VN, FP et FN - Protocole B

A partir de ces données, la sensibilité, la spécificité et l'indice de Youden de chacun de ces items ont pu être calculés (ANNEXE 7). Nous les avons regroupés dans le tableau qui suit (tableau 8).

Item n°	Protocole B				
	Se	IC(Se)95%	Spe	IC(Spe)95%	J
1. Continence labiale	0	N.C.	1	N.C.	0
2. Mvts mandibulaires	0,4	[0,10 ; 0,70]	0,86	[0,72 ; 1,01]	0,26
3. Rythme succion	0,5	[0,27 ; 0,73]	0,86	[0,67 ; 1,04]	0,36
4. Mvts linguaux	0,29	[-0,05 ; 0,62]	1	N.C.	0,29
5. Initiation déglutition pharyngée	0,08	[-0,03 ; 0,19]	1	N.C.	0,08
6. Propulsion pharyngée	0,13	[-0,04 ; 0,31]	1	N.C.	0,13
7. Continence vélopharyngée	0	N.C.	1	N.C.	0
8. Protection voies respiratoires	0	N.C.	0,96	[0,90 ; 1,03]	-0,04

N.C. : Non Calculable

Tableau 8 : Sensibilité (Se), Spécificité (Spe) et indice de Youden (J) de chaque item - Protocole B

Selon l'indice de Youden, les items les plus fiables pour ce protocole sont donc, en première place, l'item 3 « Rythme de succion-déglutition » ($J = 0,36$), puis l'item 4 « Mouvements linguaux » ($J = 0,29$) et l'item 2 « Mouvements mandibulaires » ($J = 0,26$). Se situent ensuite les items 6 « Propulsion pharyngée » ($J = 0,13$) et 5 « Initiation de la déglutition pharyngée » ($J = 0,08$), de pertinence moindre. Les items 1 « Continence labiale » et 7 « Continence vélopharyngée » présente un indice de Youden égal à 0, ce qui signifie qu'ils négligent un individu malade et un individu sain sur deux. Enfin, l'item 8 « Protection des voies respiratoires » présente un indice de Youden négatif ($J = -0,04$). Il est donc considéré comme inefficace à détecter la présence ou l'absence de trouble à ce niveau-là.

Cependant, selon notre seuil placé à 0,5, aucun de ces items n'atteste d'une bonne validité de critère.

De plus, pour le protocole B également, la spécificité est supérieure à la sensibilité, cette fois-ci pour les 8 items (voir figure 11 ci-dessous).

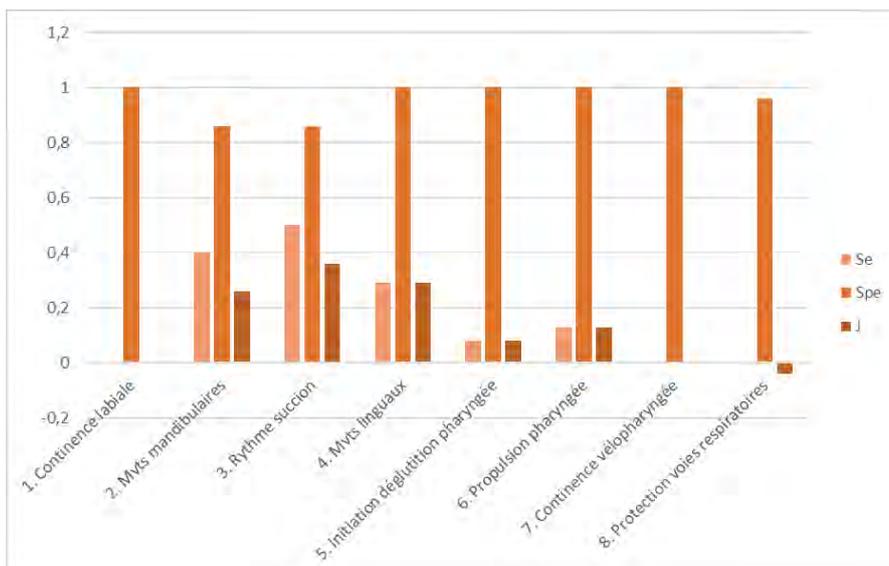


Figure 11 : Représentation de la Sensibilité (Se), de la Spécificité (Spe) et de l'indice de Youden (J) pour chaque item - Protocole B

4.2.3 Comparaison des deux protocoles

De la même manière que pour l'analyse des scores globaux, nous avons jugé pertinent de comparer les paramètres des items pour chacun des protocoles.

Pour chaque item, la spécificité du protocole B est meilleure que celle du protocole A. Cependant, la sensibilité des items du premier présente des scores bien moins élevés. En analysant parallèlement ces deux paramètres, nous pouvons ainsi constater que, pour chaque item (excepté le n°7), les résultats sont meilleurs dans le protocole A (voir figure 12 ci-dessous).

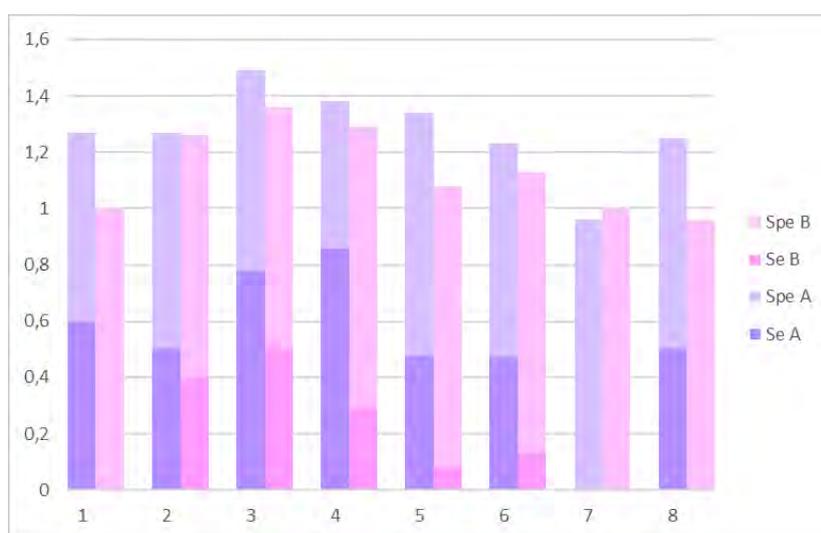


Figure 12 : Comparaison de la Sensibilité et de la Spécificité par item selon les protocoles A et B

Comme nous l'avons vu plus haut, sensibilité et spécificité analysées de manière isolée présentent un intérêt restreint quant à l'étude de la validité de critère. C'est pourquoi nous avons également jugé pertinent de comparer l'indice de Youden de chaque item, selon les deux protocoles. Cette comparaison permet de confirmer la plus grande pertinence du protocole A par rapport au B (voir figure 13 ci-dessous). Nous remarquons ici aussi que l'item 7 « Continence vélopharyngée » se détache des autres par l'insuffisance de ses résultats, quel que soit le protocole.

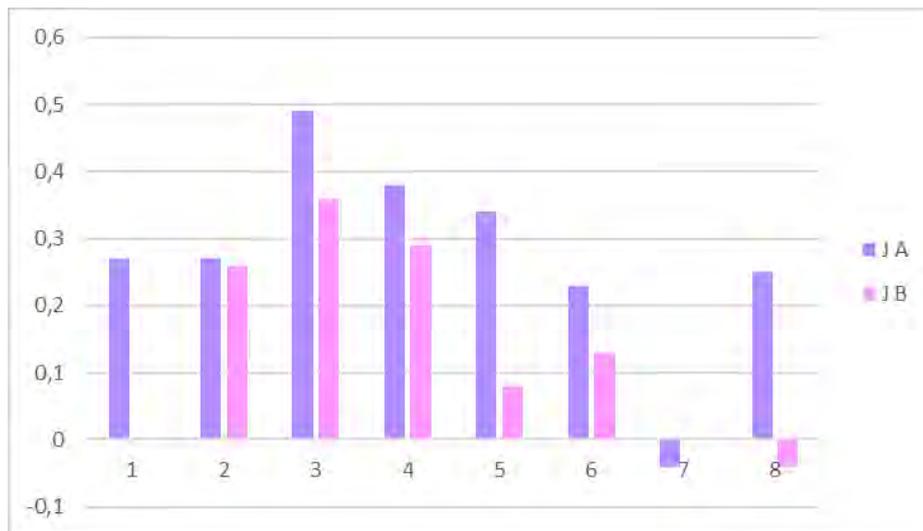


Figure 13 : Comparaison des indices de Youden pour chaque item selon les protocoles A et B

5 DISCUSSION

5.1 VERIFICATION DES HYPOTHESES

5.1.1 Hypothèse 1 : L'outil testé possède dans son ensemble une bonne validité de critère

Suite à l'étude des courbes ROC pour les protocoles A et B, nous avons obtenu les résultats suivants : $AUC(A) = 0,79$ et $AUC(B) = 0,75$. L'aire sous la courbe (AUC) était, dans les deux cas, supérieure à 0,7. Nous pouvons donc conclure que, de manière globale, cette grille d'observation clinique possède une validité de critère acceptable. Notre hypothèse 1 est ainsi validée.

5.1.2 Hypothèses 2 : Les items de l'outil testé possèdent une bonne validité de critère

Pour le protocole A comme pour le protocole B, l'analyse des items selon la valeur seuil fixée ($J \geq 0,5$) ne nous a pas permis d'attester de la bonne validité de critère d'un seul des items. En effet, l'indice de Youden de chacun des items était inférieur à cette valeur. Notre hypothèse 2 est donc invalidée.

5.2 LIMITES ET BIAIS

5.2.1 Choix de la population

Nous avons choisi, pour ce travail, de réutiliser les données de la population de l'étude OTBB3 du Pr M. Tauber. Or, le principal critère d'inclusion de cette étude était le fait d'être porteur du Syndrome de Prader-Willi. C'était donc le cas de tous les sujets de notre échantillon qui se cantonnaient ainsi à ce type de population. L'outil testé ayant cependant été créé à destination de la population générale, notre étude présentait un biais de sélection non négligeable.

De plus, la majorité des enfants porteurs de ce syndrome présentent habituellement des troubles de la déglutition et de l'alimentation. C'était le cas dans notre étude, pour laquelle 88% des sujets présentaient des troubles dysphagiques. Cette importante proportion d'enfants « malades » a sûrement déséquilibré nos calculs et ainsi, impacté nos résultats.

Si 32 comparaisons ont pu être faites, seulement 11 enfants étaient intégrés à l'étude. Chacun était vu à 3 reprises (sauf le onzième, examiné seulement deux fois). Chacun présentait donc une

possible évolution sur les 6 mois, entre la visite 1 où aucun traitement n'avait été donné et la visite 8 où tous avaient été traités, sans que nous sachions si un traitement avait été utilisé.

En effet, un autre aspect à prendre en compte, lui aussi lié directement à l'étude OTBB3, est la prise de traitement de ces enfants. L'objectif de celle-ci étant d'étudier l'effet de l'ocytocine sur les troubles de la déglutition de cette population, un certain nombre de sujets s'est vu administrer ce traitement pendant que d'autres recevaient un traitement placebo. Au vu des résultats parfois questionnants, il aurait été pertinent de savoir pour chaque enfant s'il avait bénéficié du traitement ou du placebo. L'étude était cependant menée en double aveugle, cela n'a donc pas été possible.

Nous avons de plus considéré que la prévalence du trouble était la même dans la population générale Prader-Willi (« troubles dysphagiques quasi-systématiques ») que dans notre échantillon (88%). Cependant, peut-être aurions-nous obtenu de meilleurs résultats de VPP et VPN si la prévalence de notre échantillon avait été plus élevée (par exemple > 90%). Et peut-être cette différence peut-elle être expliquée par les bénéfices des traitements prodigués.

Enfin, concernant le choix de notre population, un autre biais peut être soulevé. Si tous les nourrissons étaient âgés de moins de 3 mois au début de l'étude OTBB3, à la 8^{ème} visite (6 mois après), tous avaient plus de 6 mois. Or, la grille d'observation clinique testée avait été créée pour des nourrissons âgés de 0 à 6 mois. A V8, certains enfants commençaient à accéder au malaxage et ainsi à jouer avec la tétine, ce qui a pu impacter leur succion et donc la cotation des vidéos.

5.2.2 Procédures de passation et cotation

Les cotations de la grille d'observation clinique ont été faites à partir d'enregistrements vidéo. L'avantage est de permettre une cotation plus aisée, avec possibilité de se focaliser sur une zone, de revoir certaines séquences, de faire des arrêts sur image, etc. Cependant, ceci n'est pas représentatif d'une vraie situation d'observation de tétée. En effet, la qualité médiocre de certaines vidéos engendre une perte d'informations visuelles (pertes de lait au niveau des commissures labiales, reflux nasal, etc.) et surtout auditives (bruits de succion, respiration, stases pharyngées, ronflements, etc.).

L'enregistrement vidéo a aussi pour conséquence la perturbation de la relation parent-enfant au cours de la tétée. Certains parents pouvaient perdre de leur naturel en accentuant ou diminuant leurs interactions avec leur enfant. De plus, le nourrisson lui-même était parfois perturbé par la caméra et, dans certains cas, la suivait du regard, gênant la prise de vue et donc la cotation.

A propos de l'examen de VRDG, il est à noter que les conditions de passation n'étaient pas toujours optimales. En effet, dans certains cas, l'enfant avait bu son biberon très peu de temps avant

l'examen. Ceci a sûrement pu impacter l'appétence à la tétée et ainsi certains mécanismes de succion et de déglutition.

En outre, si, lors de l'observation de tétée, l'examineur ne devait intervenir en rien dans le positionnement et les gestes du parent, le positionnement de la tétine était systématiquement corrigé dans le cas de la VRDG, induisant ainsi une différence entre ces deux examens et donc un biais de passation.

Se pose également la question d'un potentiel biais de subjectivité de l'enquêteur. Toutes les cotations ont été effectuées par un expert du centre de référence du syndrome de Prader-Willi. Le risque est donc que les caractéristiques de ce syndrome aient été particulièrement prises en compte, aux dépens d'une cotation plus objective.

Enfin, concernant la cotation des vidéos de tétée, celle-ci a été faite à partir de la grille d'observation d'origine. C'est-à-dire que chaque item était coté sur une échelle de sévérité (normal / à risque / anormal). Or, ces données étaient ensuite analysées de manière dichotomique (normal / anormal). Si la cotation avait été binaire dès le début, l'examineur aurait sûrement pris cela en compte et les résultats en auraient peut-être été différents.

5.2.3 Sélection et appariement des items

Concernant la sélection des items, nous n'avons conservé de la grille d'observation d'origine que ceux qui pouvaient être corrélés à ceux de la grille de VRDG, ce qui représente en soi un biais.

En effet, la grille testée se veut la plus écologique possible. Un certain nombre d'items s'intéressait donc à ce qui se passe en amont de la phase orale et au contexte plus général de la tétée. Sur les 16 items, 5 ont ainsi été évincés car non objectivables lors de la VRDG : ouverture des lèvres à l'approche de la tétine, état d'éveil, comportements d'évitement, contact visuel, appréciation globale de la prise alimentaire. Ces items pourraient potentiellement être analysés en les comparant à un autre outil ciblant plus précisément ce contexte alimentaire, tel que l'OFS.

D'autres items de la grille testée n'ont pas trouvé d'items correspondants dans la grille de VRDG : contraction des muscles du plancher, tonicité des joues, rétention de la tétine par la langue.

Concernant la grille d'origine de VRDG, l'item « pauses respiratoires » n'a pas été retenu. Celui-ci ne pouvait pas être évalué de la même manière dans une situation d'observation de tétée et dans une situation d'examen de VRDG, nous avons donc jugé plus pertinent de ne pas le conserver.

5.2.4 Choix du Gold Standard

L'outil de référence que nous avons sélectionné était la VRDG, d'une part pour des raisons pratiques : nous avons à notre disposition les vidéos de cet examen des enfants de notre échantillon. D'autre part, c'était un examen reconnu par les professionnels de santé de la petite enfance, qui représentait donc un moyen sûr de juger la présence ou l'absence de troubles dysphagiques.

Cependant, la VRDG et l'observation de tétée présentent des situations d'examen très différentes, tant au niveau du contexte de passation que de l'environnement et de l'ambiance dans lesquels ils se déroulent. Si l'un peut avoir lieu de manière presque écologique, en présence des parents, avec les instruments d'alimentation habituels, l'autre, en revanche, est une réelle situation d'examen pendant laquelle l'enfant se retrouve seul, séparé de ses parents par une vitre, dans un contexte très différent d'un temps de repas ordinaire. Cela peut donc engendrer du stress, ou du moins des variations dans les mécanismes de déglutition. Il est également à noter que l'enfant n'est évalué qu'à partir de la mise en bouche, ce qui a pour conséquence, nous l'avons vu, l'éviction de plusieurs items de la grille dans notre étude.

Il convient donc de replacer les résultats statistiques dans ce contexte (notamment pour l'analyse par item) et de se questionner sur l'impact qu'a pu avoir cette inégalité de contextes dans l'analyse statistique. Peut-être est-il pertinent de se demander si la validité de critère peut-être rigoureusement obtenue en comparant des examens de forme si différente.

En effet, le mieux aurait sûrement été de comparer notre outil testé à une autre grille déjà existante. Cependant comme nous l'avons vu dans la partie théorique, aucune grille suffisamment validée n'aurait pu servir d'outil de référence, d'où la création de celle-ci.

5.3 RETOUR D'EXPERIENCE ET APPLICATIONS FUTURES

L'analyse des score globaux nous a permis de conclure à une bonne validité de critère de la grille testée. Afin de permettre une meilleure interprétation des résultats obtenus et de déterminer les applications futures de cette grille, nous avons cependant décidé d'analyser de manière plus fine et de mettre en lien les résultats obtenus.

5.3.1 Retour sur les résultats des scores globaux

Dans le protocole A comme dans le protocole B, les résultats des scores globaux témoignaient d'une meilleure spécificité par rapport à la sensibilité. Ainsi, cette grille possède une meilleure capacité

à identifier l'absence d'un trouble chez un individu. Autrement dit, lorsqu'un individu n'est porteur d'aucun trouble dysphagique, la grille ne l'identifie jamais comme étant « malade », mais comme « sain » de manière fiable, quitte à accepter dans cette catégorie un certain nombre de « malades » (donc de FN) (voir tableaux 9 et 10 ci-dessous). Nous pouvons en conclure que cette grille est plus fiable pour les patients « sains ». Cela correspond ainsi à un test de type diagnostic davantage qu'à un test de type dépistage.

Ces résultats sont corroborés par les valeurs de VPP observées. En effet, pour le protocole A comme pour le protocole B, on observe de meilleures valeurs pour la VPP que pour la VPN. Concrètement, si un test est positif, on est sûr que l'individu testé est malade (meilleure VPP) alors que si un test est négatif, le caractère sain ou malade de la personne testée est moins certain (il peut être sain ou malade). Ainsi, la grille permet de confirmer un trouble dysphagique, quitte à ne pas repérer certains individus malades (diagnostic), plutôt que de recenser tous les individus qui peuvent être porteurs du trouble, quitte à en identifier en trop (dépistage) (voir tableaux 9 et 10 ci-dessous).

S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M

Légende

S : sain

M : malade

□ : "sain" selon l'outil testé

□ : "malade" selon l'outil testé

Tableau 9 : Représentation concrète des 32 vidéos en fonction de leur statut normal/anormal et de leur résultat à la grille d'observation - Protocole A

S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M
S	M	M	M	M	M	M	M

Légende

S : sain

M : malade

□ : "sain" selon l'outil testé

□ : "malade" selon l'outil testé

Tableau 10 : Représentation concrète des 32 vidéos en fonction de leur statut normal/anormal et de leur résultat à la grille d'observation - Protocole B

Ainsi, ces 2 protocoles correspondent à des tests de type diagnostic. On peut cependant introduire une nuance entre eux : avec le protocole A, la proportion d'individus malades avec un test négatif (score en dessous du seuil) est d'environ $1/3$ (10 patients malades mais considérés « sains » sur 32), tandis qu'avec le protocole B, cette proportion s'élève à quasiment $1/2$ (14 patients malades mais considérés « sains » sur 32) (voir tableaux 9 et 10 ci-dessus). Le protocole B aura donc plus facilement tendance à laisser passer des individus malades avec un résultat de test négatif. Par conséquent, si l'on devait choisir entre les 2 protocoles à partir de ces considérations, il serait donc plus judicieux de s'orienter vers le protocole A.

Une deuxième nuance est à prendre en considération dans la distinction de ces deux protocoles. En effet, si les résultats obtenus par le protocole B avaient été pertinents, la grille aurait eu la capacité, de par sa distinction normal/à risque VS anormal, de repérer uniquement les troubles de type sévère. Le protocole A, quant à lui, permet donc un diagnostic efficace des sujets présentant des troubles de la déglutition, qu'ils soient peu ou très sévères, sans distinction du degré de sévérité.

Nous pouvons donc imaginer, comme application future de cette grille d'observation clinique, une utilisation dichotomique de type normal/anormal, plutôt qu'une échelle de sévérité en trois critères comme le proposait la grille d'origine. Cette idée semble appuyée par les résultats obtenus dans le mémoire d'Emilie Leboucher. En effet, la fidélité inter-juge de cette grille a été testée à partir de la grille d'origine proposant 3 niveaux de sévérité. Or, les résultats obtenus n'ayant pas été concluants, nous nous questionnons d'autant plus quant à la pertinence de la cotation de cette grille de manière binaire.

5.3.2 Retour sur les résultats par item

Le manque de significativité des résultats par item ne permettait pas de conclure à une bonne validité de critère de ceux-ci.

Les indices de Youden peu élevés s'expliquent par les valeurs de la sensibilité et de la spécificité à partir desquelles ils étaient calculés. En effet, soit les valeurs de ces dernières étaient trop peu élevées, soit sensibilité et spécificité du même item étaient trop éloignées l'une de l'autre, affaiblissant ainsi l'indice de Youden. De plus, les intervalles de confiance à 95% affichaient régulièrement, lorsqu'ils étaient calculables, des bornes éloignées l'une de l'autre. Ceci témoigne d'une incertitude non négligeable quant à la fiabilité des valeurs de la sensibilité et de la spécificité, et donc par extension, de l'indice de Youden.

Cependant, pour 6 items sur 8 du protocole A, l'indice de Youden sur lequel nous nous basions était supérieur à 0,25. Nous nous sommes donc questionné sur la pertinence de notre choix de valeur seuil. En l'abaissant, notre hypothèse 2 aurait été validée, mais la fiabilité des items en aurait sûrement été impactée.

Pour le protocole B en revanche, même en abaissant ce seuil à 0,25, seuls 3 items sur les 8 auraient été valides, les indices de Youden étant tous globalement plus bas du fait des sensibilités et spécificités très hétérogènes.

Ainsi, malgré ces résultats, la validité de critère globale de la grille a pu être démontrée et nous avons pu constater que certains items se démarquaient plus particulièrement des autres. Par exemple, les items 3 (« Rythme de succion », $J(A) = 0,49$ et $J(B) = 0,36$) et 4 (« Mouvements linguaux », $J(A) = 0,38$ et $J(B) = 0,29$) sont ceux qui obtiennent le meilleur indice de Youden, dans le protocole A comme dans le B. Le premier protocole privilégie ensuite l'item 5 (« Initiation de la déglutition pharyngée », $J(A) = 0,34$), tandis que le second met en avant le 2 (« Mouvements mandibulaires », $J(B) = 0,26$).

Ainsi, peut-être ces 4 items seraient-ils les plus spécifiques et pertinents à observer et donc ceux à privilégier dans une démarche diagnostique de troubles de la déglutition. Se concentrer sur ceux-ci pourrait permettre l'élaboration d'un outil plus accessible et dont la prise en main serait plus aisée, surtout si celui-ci est à destination de personnes proches de l'entourage ou de professionnels non formés.

5.3.3 Ebauche de validation de la sensibilité au changement

Bien que toutes les conditions ne soient pas réunies pour permettre une analyse rigoureuse de la sensibilité au changement de la grille d'observation, nous avons cependant pu en commencer l'ébauche de manière qualitative.

La sensibilité au changement est à la capacité d'un instrument à percevoir avec justesse une évolution dans les aptitudes mesurées (Fermanian, 2005).

Nous étudions dans notre étude 11 patients examinés chacun à 3 reprises (sauf le dernier). Aucun n'avait de traitement à la première visite et tous en avait un, ocytocine ou placebo, à la dernière. En nous appuyant sur les scores obtenus à chaque vidéo, patient par patient (voir figures 1 et 3), nous avons pu avoir un aperçu de ce paramètre.

Pour le protocole A, 5 patients sur les 11 ont vu leur score diminuer entre V1 et V8. Pour le protocole B, 8 d'entre eux ont également témoigné d'une amélioration de par la diminution de leur

score à la grille d'observation clinique. Cependant, les nombreux biais (notamment concernant le traitement par ocytocine dont l'efficacité n'est pas encore prouvée) et le peu de rigueur de ces observations ne permettent pas de conclure concernant la sensibilité au changement. Une étude plus approfondie de ce paramètre serait ainsi nécessaire afin de continuer la validation de cet outil.

5.4 PERSPECTIVES

Au vu des résultats obtenus, des ajustements peuvent être apportés à notre grille d'observation clinique, afin d'en faire un outil valide et facile d'utilisation, à destination de tout professionnel de la petite enfance, en salariat ou en libéral.

Cet outil pourrait être retravaillé en remplaçant l'échelle de sévérité à trois degrés en une cotation binaire de type normal/anormal. De plus, l'inclusion d'une échelle de sévérité (par exemple, de 1 à 5) sur l'ensemble de la tétée permettrait une analyse du ressenti global tout en mettant en évidence les items les plus représentatifs d'un haut niveau de sévérité. Celui-ci serait ainsi plus accessible et resterait tout aussi pertinent dans son objectif diagnostique.

Suite aux améliorations apportées, il serait également intéressant de réévaluer la validité de critère item par item, que nous n'avons pu démontrer dans cette étude. Avec une cotation binaire et un échantillon plus grand et représentatif de la population générale, nous pouvons espérer obtenir des résultats plus significatifs. Il serait également intéressant de mettre en évidence les items qui se sont révélés, au cours de cette étude, les plus évocateurs d'un trouble dysphagique.

Une vérification de la validité de critère des items que nous n'avons pu tester (concernant l'alimentation et le contexte de repas) serait aussi intéressante pour compléter notre étude. La comparaison avec une évaluation comportementale validée apporterait cet éclairage supplémentaire, grâce par exemple aux grilles d'observation de Chatoor (Chatoor et al., 2018), ou à l'OFS.

Enfin, la sensibilité au changement, que nous n'avons pu qu'ébaucher dans cette étude, apporterait une propriété psychométrique supplémentaire à cet outil. Il serait intéressant de l'analyser plus rigoureusement, cette fois-ci de manière quantitative.

6 CONCLUSION

Face à l'augmentation de l'incidence des troubles de l'alimentation et de la déglutition dans la population pédiatrique et au manque d'outils adaptés et validés rigoureusement, la création et la validation de nouveaux tests devient une priorité.

Nous cherchions donc, dans ce mémoire, à poursuivre la validation de la grille d'observation clinique précédemment élaborée et partiellement validée par les mémoires de Pauline Crusson (2013), Morgane Péron (2018) et Camille Poumarède (2019). L'étude de la validité de critère de cette grille a ainsi permis de conclure à une bonne capacité globale de celle-ci à identifier fidèlement la présence ou l'absence de troubles dysphagiques chez le nourrisson.

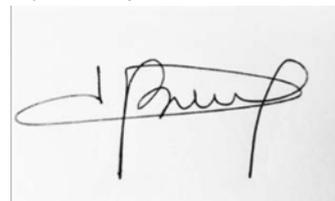
De plus, les bonnes capacités diagnostiques de cet outil permettront, à l'avenir, d'éviter un examen de VRDG, coûteux et à risque pour l'enfant, lorsque celui-ci n'est pas nécessaire.

Enfin, l'adaptation binaire du système de cotation pourra permettre une utilisation plus aisée de l'outil. Tout professionnel de santé de la petite enfance y aurait ainsi facilement accès, dont les orthophonistes en libéral qui se retrouvent parfois démunies face aux difficultés de l'alimentation et/ou de la déglutition de patients aussi jeunes.

Cette étude pourrait être complétée et enrichie, notamment par le retravail du système de cotation, par l'élargissement de la population testée dans l'analyse de la validité de critère par item, par l'apport de nouvelles données sur le contexte alimentaire et le comportement du nourrisson (grâce à l'OFS ou aux grilles de Chatoor) ou encore par la poursuite de la standardisation de l'outil (avec l'étude de sa sensibilité au changement).



Dr Pascale FICHAUC BOURIN
Praticien Hospitalier
Service d'ORL et Chirurgie cervico-faciale
Hôpital Larrey TOULOUSE



BIBLIOGRAPHIE

- Arvedson, J. C. (2000). Evaluation of Children With Feeding and Swallowing Problems. *Language, Speech, and Hearing Services in Schools, 31*(1), 28-41. <https://doi.org/10.1044/0161-1461.3101.28>
- Arvedson, J. C. (2008). Assessment of pediatric dysphagia and feeding disorders : Clinical and instrumental approaches. *Developmental Disabilities Research Reviews, 14*(2), 118-127. <https://doi.org/10.1002/ddrr.17>
- Arvedson, J. C., Brodsky, L., & Lefton-Greif, M. A. (2019). *Pediatric Swallowing and Feeding : Assessment and Management, Third Edition*. Plural Publishing.
- Arvedson, J. C., & Lefton-Greif, M. A. (2017). Instrumental Assessment of Pediatric Dysphagia. *Seminars in Speech and Language, 38*(2), 135-146. <https://doi.org/10.1055/s-0037-1599111>
- Bickell, M., Barton, C., Dow, K., & Fucile, S. (2017). A systematic review of clinical and psychometric properties of infant oral motor feeding assessments. *Developmental Neurorehabilitation, 1-11*. <https://doi.org/10.1080/17518423.2017.1289272>
- Brodsky, M. B., Mayfield, E. B., & Gross, R. D. (2019). Clinical Decision Making in the ICU : Dysphagia Screening, Assessment, and Treatment. *Seminars in Speech and Language, 40*(3), 170-187. <https://doi.org/10.1055/s-0039-1688980>
- Cascales, T. (2015). *Les troubles alimentaires du bébé : Approche psychanalytique et développementale*. Éditions Érès.
- Centre Référent SPW. (2020). *Procédure Radioscopie de déglutition—Protocole OTBB3*.
- Chatoor, I., Hommel, S., Sechi, C., & Lucarelli, L. (2018). Development of the parent-child play scale for use in children with feeding disorders : Parent-Child Play Scale in Early Feeding Disorders. *Infant Mental Health Journal, 39*(2), 153-169. <https://doi.org/10.1002/imhj.21702>

- Crusson, P. P. (2013). *Etude préliminaire d'une grille d'observation des fonctions oro-myo-fonctionnelles du nourrisson âgé de 1 à 6 mois, alimenté au biberon.*
- Delacour, H., Servonnet, A., Perrot, A., Vigezzi, J. F., & Ramirez, J. M. (2005). [ROC (receiver operating characteristics) curve : Principles and application in biology]. *Annales De Biologie Clinique*, 63(2), 145-154.
- DeMatteo, C., Matovich, D., & Hjartarson, A. (2005). Comparison of clinical and videofluoroscopic evaluation of children with feeding and swallowing difficulties. *Developmental Medicine & Child Neurology*, 47(3), 149-157. <https://doi.org/10.1017/S0012162205000289>
- Dodrill, P., & Gosa, M. M. (2015). Pediatric Dysphagia : Physiology, Assessment, and Management. *Annals of Nutrition and Metabolism*, 66(Suppl. 5), 24-31. <https://doi.org/10.1159/000381372>
- Duffy, K. L. (2018). Dysphagia in Children. *Current Problems in Pediatric and Adolescent Health Care*, 48(3), 71-73. <https://doi.org/10.1016/j.cppeds.2018.01.003>
- Fermanian, J. (2005). Validation des échelles d'évaluation en médecine physique et de réadaptation : Comment apprécier correctement leurs qualités psychométriques. *Annales de Réadaptation et de Médecine Physique*, 48(6), 281-287. <https://doi.org/10.1016/j.annrmp.2005.04.004>
- Goday, P. S., Huh, S. Y., Silverman, A., Lukens, C. T., Dodrill, P., Cohen, S. S., Delaney, A. L., Feuling, M. B., Noel, R. J., Gisel, E., Kenzer, A., Kessler, D. B., Kraus de Camargo, O., Browne, J., & Phalen, J. A. (2019). Pediatric Feeding Disorder : Consensus Definition and Conceptual Framework. *Journal of Pediatric Gastroenterology and Nutrition*, 68(1), 124-129. <https://doi.org/10.1097/MPG.0000000000002188>
- Heckathorn, D.-E., Speyer, R., Taylor, J., & Cordier, R. (2016). Systematic Review : Non-Instrumental Swallowing and Feeding Assessments in Pediatrics. *Dysphagia*, 31(1), 1-23. <https://doi.org/10.1007/s00455-015-9667-5>

- Howe, T.-H., Lin, K.-C., Fu, C.-P., Su, C.-T., & Hsieh, C.-L. (2008). A Review of Psychometric Properties of Feeding Assessment Tools Used in Neonates. *Journal of Obstetric, Gynecologic & Neonatal Nursing, 37*(3), 338-349. <https://doi.org/10.1111/j.1552-6909.2008.00240.x>
- Kleinert, J. Or. (2017). Pediatric Feeding Disorders and Severe Developmental Disabilities. *Seminars in Speech and Language, 38*(2), 116-125. <https://doi.org/10.1055/s-0037-1599109>
- Kumar, S. P., Mooney, R., Wieser, L. J., & Havstad, S. (2006). The LATCH Scoring System and Prediction of Breastfeeding Duration. *Journal of Human Lactation, 22*(4), 391-397. <https://doi.org/10.1177/0890334406293161>
- Lawlor, C. M., & Choi, S. (2020). Diagnosis and Management of Pediatric Dysphagia : A Review. *JAMA Otolaryngology-- Head & Neck Surgery, 146*(2), 183-191. <https://doi.org/10.1001/jamaoto.2019.3622>
- Lefton-Greif, M. A. (2008). Pediatric Dysphagia. *Physical Medicine and Rehabilitation Clinics of North America, 19*(4), 837-851. <https://doi.org/10.1016/j.pmr.2008.05.007>
- Lefton-Greif, M. A., & Arvedson, J. C. (2016). Pediatric Feeding/Swallowing : Yesterday, Today, and Tomorrow. *Seminars in Speech and Language, 37*(4), 298-309. <https://doi.org/10.1055/s-0036-1587702>
- Longoni, L., Provenzi, L., Cavallini, A., Sacchi, D., Scotto di Minico, G., & Borgatti, R. (2018). Predictors and outcomes of the Neonatal Oral Motor Assessment Scale (NOMAS) performance : A systematic review. *European Journal of Pediatrics, 177*(5), 665-673. <https://doi.org/10.1007/s00431-018-3130-1>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & 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>

Péron, M. (2018). *Élaboration et prévalidation d'une grille d'évaluation non instrumentale des capacités d'alimentation et de déglutition du nourrisson alimenté au biberon.*

file:///C:/Users/charl/Documents/Orthophonie/CFUO%20Toulouse/5%C3%A8me%20ann%C3%A9e/M%C3%A9moire/Documents%20M%C3%A9moire%20Allaitement/m%C3%A9moire%20morgane%20p%C3%A9ron.pdf

Poumarède, C. (2019). *Validation de la fidélité d'un outil d'évaluation non instrumentale des capacités d'alimentation et de déglutition du nourrisson alimenté au biberon.*

Protocole National de Diagnostic et de Soins (PNDS)—Syndrome Prader-Willi. (2021).

Tauber, M., Boulanouar, K., Diene, G., Çabal-Berthoumieu, S., Ehlinger, V., Fichaux-Bourin, P., Molinas, C., Faye, S., Valette, M., Pourrinet, J., Cessans, C., Viaux-Sauvelon, S., Bascoul, C., Guedeney, A., Delhanty, P., Geenen, V., Martens, H., Muscatelli, F., Cohen, D., ... Salles, J.-P. (2017). The Use of Oxytocin to Improve Feeding and Social Skills in Infants With Prader–Willi Syndrome. *Pediatrics*, 139(2), e20162976. <https://doi.org/10.1542/peds.2016-2976>

Tauber, P. M., & Fichaux-Bourin, D. P. (s. d.). *Présentation de l'étude—Traitement par ocytocine des nourrissons présentant un syndrome de Prader-Willi : Effets d'administrations intranasales d'ocytocine chez des nourrissons âgés de moins de 3 mois vs. Placebo sur le comportement alimentaire.*

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & 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>

Tétréault, S., & Guillez, P. (2014). *Guide pratique de recherche en réadaptation.* De Boeck - Solal.

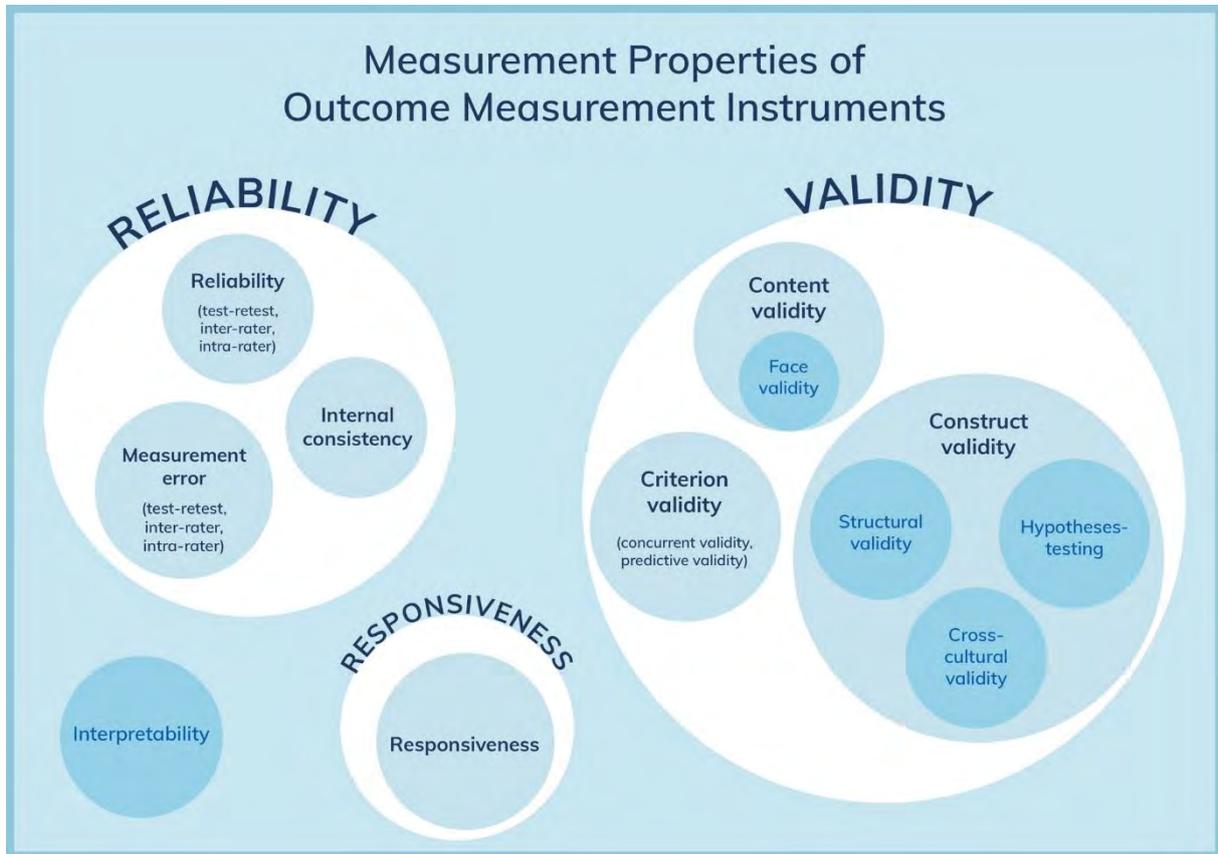
Zarem, C., Kidokoro, H., Neil, J., Wallendorf, M., Inder, T., & Pineda, R. (2013). Psychometrics of the Neonatal Oral Motor Assessment Scale. *Developmental Medicine & Child Neurology*, 55(12), 1115-1120. <https://doi.org/10.1111/dmcn.12202>

TABLE DES ANNEXES

ANNEXE 1 : The COSMIN Taxonomy	61
ANNEXE 2 : Grille d'observation clinique.....	62
ANNEXE 3 : Grille de cotation de la vidéoradioscopie de la déglutition.....	63
ANNEXE 4 : Tableau d'appariement des deux grilles de cotation.....	64
ANNEXE 5 : Tableaux de résultats des scores globaux - Protocole A et B	65
ANNEXE 6 : Représentation du nombre de patients malades (M+) et sains (M-) en fonction du score de la grille d'observation – Protocoles A et B	66
ANNEXE 7 : Tableaux de contingence et calculs statistiques par item - Protocoles A et B.....	67

ANNEXES

ANNEXE 1 : The COSMIN Taxonomy



Source : The COSMIN checklist

<https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/>

ANNEXE 2 : Grille d'observation clinique

	NORMAL (1)	A RISQUE (2)	ANORMAL (3)	RESULTATS (1 / 2 / 3)
Item 1 : Continence labiale	Continence labiale 1- Maintenu	2- Non maintenue	3- Absente	
Item 2 : Mouvements mandibulaires	Mouvements mandibulaires adaptés 1 - Mouvements adaptés	2 - Mouvements trop amples	3 - Mouvements inefficaces	
Item 3 : Coordination des mvts - Rythme de succion	Rythme de succion-déglutition pdt 2 min 1 - > 10 succions-déglutitions	2 - < 10 succions-déglutitions	3 - Pas de rythme	
Item 4 : Synchronisation Succion-Déglutition	Synchronisation Succion-Déglutition 1 - Bonne synchronisation S-D (1 à 3)	2 - Désynchro S-D puis retour	3 - Majoration désynchronisation S-D	
Item 5 : Synchronisation Déglutition-Respiration	Synchronisation Déglutition-Respiration 1 - Bonne synchronisation D-R	2 - Désynchro D-R puis récupération	3 - Désynchronisation D-R sans récup	
Item 6 : Propulsion pharyngée	Bruits respiratoires normaux 1 - Petits bruits respiratoires libres	2 - Bruits pharyngés intermittents	3 - Majoration encombrement respi	
Item 7 : Continence vélopharyngée	Écoulement nasal de lait 1 - Absence de reflux	2 - Léger reflux nasal	3 - Reflux nasal important	
Item 8 : Protection des voies respiratoires	Mécanismes d'expulsion 1 - Aucun signe ou justifié	2 - Présence, récurrents	3 - Aucun mécanisme malgré FR	

ANNEXE 3 : Grille de cotation de la vidéoradioscopie de la déglutition

N° de protocole	Acronyme de l'étude	Patient __ __
RC/31/15/7825	OTBB3	Visite __ __

Promoteur CHU TOULOUSE

Patient number : |__| - |__| Visit |__| Date |__|/|__|/|__|

Grille Cotation Radioscopie

Domaine	Composante	Observation	Résultats Normal (0 pt) / Anormal (1 pt)
Phase orale	Fermeture labiale	Continente	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)
	Mouvements mandibulaires	Stabilité	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)
	Mouvements linguaux	Nombre de mouvements de succion par déglutition	<input type="checkbox"/> 1 à 3 (0pt) <input type="checkbox"/> >3 ou 0 (1pt)
	Coordination des mouvements	Rythme de succion régulier	<input type="checkbox"/> Oui (0pt) <input type="checkbox"/> Non (1pt)
Initiation et synchronisation de la déglutition	Initiation de la déglutition pharyngée	Localisation du bolus	<input type="checkbox"/> Base de langue (0pt) <input type="checkbox"/> Aux vallécules (1pt) <input type="checkbox"/> Sinus piriformes (1pt)
Phase pharyngée et protection des voies aériennes	Contenance vélopharyngée	Reflux nasopharyngé	<input type="checkbox"/> Non (0pt) <input type="checkbox"/> Oui (1pt)
	Propulsion pharyngée	Présence de stases	<input type="checkbox"/> Non (0pt) <input type="checkbox"/> Oui (1pt) ° entre la base de langue et les vallécules ° Après les vallécules (sinus piriformes / murs pharyngés)
	Protection des voies aériennes	Pénétrations et inhalations	<input type="checkbox"/> Non (0pt) <input type="checkbox"/> Oui (1pt) <input type="checkbox"/> Pénétration expulsée <input type="checkbox"/> Inhalation expulsée <input type="checkbox"/> Pénétration non-expulsée <input type="checkbox"/> Inhalation non- expulsée

ANNEXE 4 : Tableau d'appariement des deux grilles de cotation

		Grille Observation Clinique				Grille VRDG			
DOMAINE		ANALYSE SCORES GLOBAUX		ANALYSE PAR ITEM		ANALYSE SCORES GLOBAUX	RESULTAT 0 pt / 1 pt	RESULTAT M- / M+	
		RESULTAT 0 pt (1) / 1 pt (2-3) Protocole A	RESULTAT 1-2 (0 pt) / 3 (1 pt) Protocole B	RESULTAT T- (1) / T+ (2-3) Protocole A	RESULTAT T- (1-2) / T+ (3) Protocole B				
		NORMAL (0 pt ou -) / ANORMAL (1 pt ou +)				NORMAL (0 pt ou -) / ANORMAL (1 pt ou +)			
PHASE ORALE	Item 1 : Continence labiale	Continence labiale 1- Maintenu (0) 2- Non maintenue (0/1) 3- Absente (1)				Continence labiale Oui (0 pt) Non (1 pt)			
	Item 2 : MVts mandibulaires	MVts mandibulaires adaptés 1- Mouvements adaptés (0) 2- Mouvements trop amples (0/1) 3- Mouvements inefficaces (1)				MVts stables Oui (0 pt) Non (1 pt)			
	Item 3 : Rythme de succion	Rythme de succion-déglutition pdt 2 min. 1- > 10 succions-déglutitions (0) 2- < 10 succions-déglutitions (0/1) 3- Pas de rythme (1)				Rythme de succion régulier Oui (0 pt) Non (1 pt)			
	Item 4 : Mouvements linguaux	Synchronisation Succion-Déglutition 1- Bonne synchronisation S-D (1 à 3) (0) 2- Désynchro S-D puis retour (0/1) 3- Majoration désynchronisation S-D (1)				Nb de mvts de succion par déglutition 1 à 3 (0 pt) > 4 ou 0 (1 pt)			
	Item 5 : Initiation déglutition pharyngée	Synchronisation Déglutition-Respiration 1- Bonne synchronisation D-R (0) 2- Désynchro D-R puis récup (0/1) 3- Désynchronisation D-R sans récup (1)				Localisation du bolus Base de langue (0 pt) Aux vallécules (1 pt) Sinus piriformes (1 pt)			
	Item 6 : Propulsion pharyngée	Bruits respiratoires normaux 1- Petits bruits respiratoires libres (0) 2- Bruits pharyngés intermittents (0/1) 3- Majoration encombrement respi (1)				Présence de stases Non (0 pt) Oui (1 pt) * Entre Bdt et vallécules * Après vallécules (sinus piriformes / mus pharyngés)			
	Item 7 : Continence vélopharyngée	Ecoulement nasal de lait 1- Absence de reflux (0) 2- Léger reflux nasal (0/1) 3- Reflux nasal important (1)				Reflux nasopharyngé Non (0 pt) Oui (1 pt)			
	Item 8 : Protection des voies respiratoires	Mécanismes d'expulsion 1- Aucun signe ou justifié (0) 2- Présence, récurrents (0/1) 3- Aucun mécanisme malgré FR (1)				Pénétrations ou inhalations Non (0 pt) Oui (1 pt) - Pénétration expulsee - Inhalation expulsee - Pénétration non-expulsee - Inhalation non-expulsee			
	TOTAUX	0	0	0	0	0	0		
		/8	/8	/8	/8	/8	/8		

ANNEXE 5 : Tableaux de résultats des scores globaux - Protocole A et B

Protocole A

N° Anonymat	Appariement N°	Score /8 Grille tétée <i>Protocole A</i>	Patho (selon VR) (M- : 0 / M+ : 1)	Se	Spe	1-Spe	Youden Se + Spe-1
01_03_V8	9	0	1	1	0	1	0
01_04_V8	12	0	0	1	0	1	0
01_07_V4	20	0	1	1	0	1	0
01_08_V1	22	0	1	1	0	1	0
01_08_V4	23	0	1	1	0	1	0
01_10_V4	29	0	1	1	0	1	0
01_01_V4	2	1	1	0,82142857	0,25	0,75	0,07142857
01_01_V8	3	1	0	0,82142857	0,25	0,75	0,07142857
01_03_V4	8	2	1	0,78571429	0,5	0,5	0,28571429
01_06_V1	16	2	1	0,78571429	0,5	0,5	0,28571429
01_06_V8	18	2	0	0,78571429	0,5	0,5	0,28571429
01_07_V8	21	2	1	0,78571429	0,5	0,5	0,28571429
01_08_V8	24	2	0	0,78571429	0,5	0,5	0,28571429
01_09_V1	25	2	1	0,78571429	0,5	0,5	0,28571429
01_03_V1	7	3	1	0,64285714	1	0	0,64285714
01_04_V4	11	3	1	0,64285714	1	0	0,64285714
01_05_V1	13	3	1	0,64285714	1	0	0,64285714
01_05_V8	15	3	1	0,64285714	1	0	0,64285714
01_06_V4	17	3	1	0,64285714	1	0	0,64285714
01_02_V8	6	4	1	0,46428571	1	0	0,46428571
01_07_V1	19	4	1	0,46428571	1	0	0,46428571
01_09_V4	26	4	1	0,46428571	1	0	0,46428571
01_10_V1	28	4	1	0,46428571	1	0	0,46428571
01_10_V8	30	4	1	0,46428571	1	0	0,46428571
01_11_V1	31	4	1	0,46428571	1	0	0,46428571
01_02_V4	5	5	1	0,25	1	0	0,25
01_04_V1	10	5	1	0,25	1	0	0,25
01_05_V4	14	5	1	0,25	1	0	0,25
01_01_V1	1	6	1	0,14285714	1	0	0,14285714
01_09_V8	27	6	1	0,14285714	1	0	0,14285714
01_11_V8	32	6	1	0,14285714	1	0	0,14285714
01_02_V1	4	7	1	0,03571429	1	0	0,03571429

= Valeur seuil optimale

- 01_01 = Patient 01
- 01_02 = Patient 02
- 01_03 = Patient 03
- 01_04 = Patient 04
- 01_05 = Patient 05
- 01_06 = Patient 06
- 01_07 = Patient 07
- 01_08 = Patient 08
- 01_09 = Patient 09
- 01_10 = Patient 10
- 01_11 = Patient 11

Protocole B

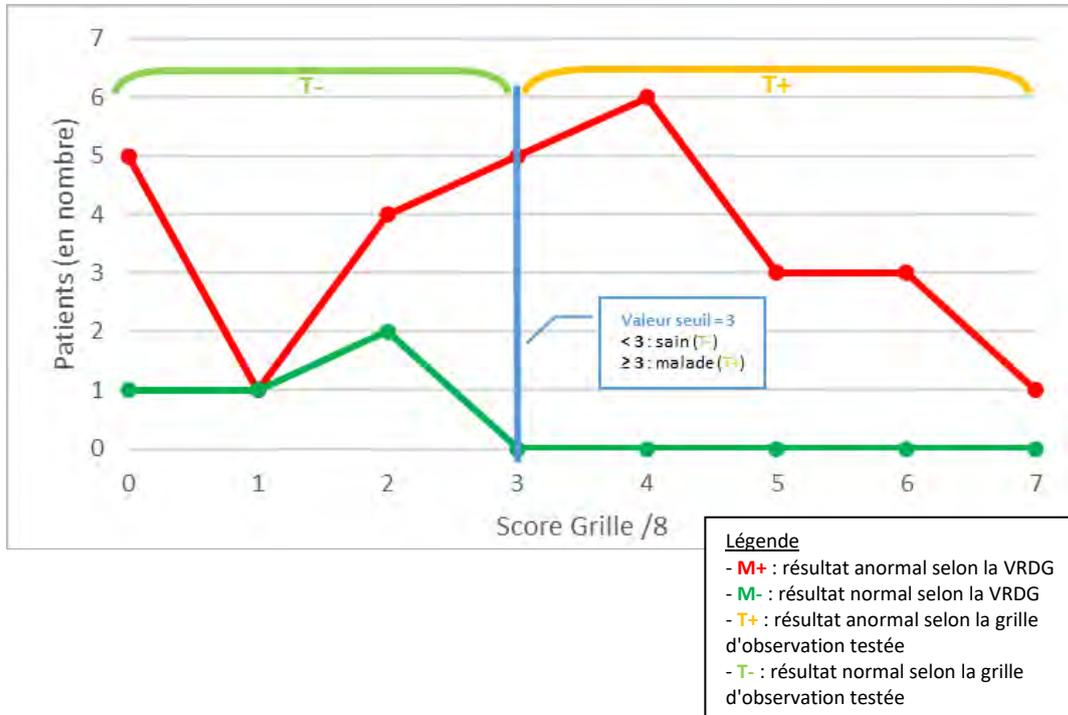
N° Anonymat	Appariement N°	Score /8 Grille tétée <i>Protocole B</i>	Patho (selon VR) (M- : 0 / M+ : 1)	Se	Spe	1-Spe	Youden Se + Spe-1
01_01_V4	2	0	1	1	0	1	0
01_01_V8	3	0	0	1	0	1	0
01_02_V4	5	0	1	1	0	1	0
01_02_V8	6	0	1	1	0	1	0
01_03_V4	8	0	1	1	0	1	0
01_03_V8	9	0	1	1	0	1	0
01_04_V4	11	0	1	1	0	1	0
01_04_V8	12	0	0	1	0	1	0
01_05_V8	15	0	1	1	0	1	0
01_06_V8	18	0	0	1	0	1	0
01_07_V1	19	0	1	1	0	1	0
01_07_V4	20	0	1	1	0	1	0
01_07_V8	21	0	1	1	0	1	0
01_08_V1	22	0	1	1	0	1	0
01_08_V4	23	0	1	1	0	1	0
01_08_V8	24	0	0	1	0	1	0
01_10_V4	29	0	1	1	0	1	0
01_10_V8	30	0	1	1	0	1	0
01_02_V1	4	1	1	0,5	1	0	0,5
01_03_V1	7	1	1	0,5	1	0	0,5
01_05_V4	14	1	1	0,5	1	0	0,5
01_06_V4	17	1	1	0,5	1	0	0,5
01_09_V1	25	1	1	0,5	1	0	0,5
01_09_V8	27	1	1	0,5	1	0	0,5
01_11_V8	32	1	1	0,5	1	0	0,5
01_01_V1	1	2	1	0,25	1	0	0,25
01_04_V1	10	2	1	0,25	1	0	0,25
01_05_V1	13	2	1	0,25	1	0	0,25
01_06_V1	16	2	1	0,25	1	0	0,25
01_09_V4	26	3	1	0,10714286	1	0	0,10714286
01_10_V1	28	3	1	0,10714286	1	0	0,10714286
01_11_V1	31	3	1	0,10714286	1	0	0,10714286

= Valeur seuil optimale

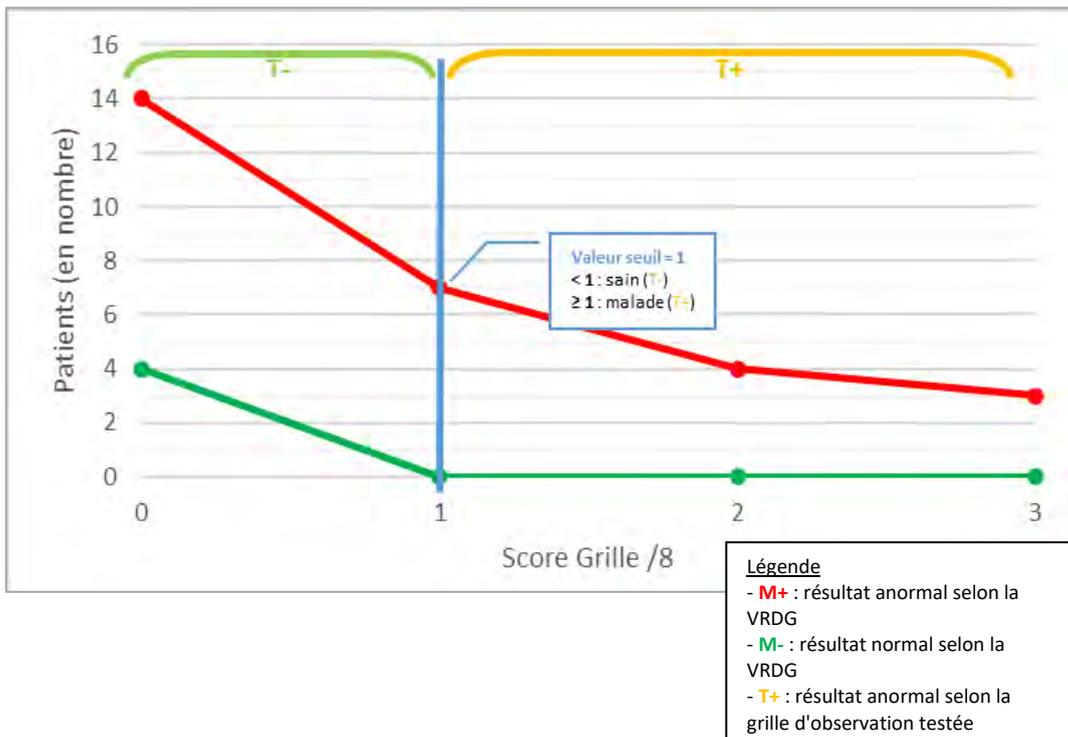
- 01_01 = Patient 01
- 01_02 = Patient 02
- 01_03 = Patient 03
- 01_04 = Patient 04
- 01_05 = Patient 05
- 01_06 = Patient 06
- 01_07 = Patient 07
- 01_08 = Patient 08
- 01_09 = Patient 09
- 01_10 = Patient 10
- 01_11 = Patient 11

ANNEXE 6 : Représentation du nombre de patients malades (M+) et sains (M-) en fonction du score de la grille d'observation – Protocoles A et B

Protocole A



Protocole B



ANNEXE 7 : Tableaux de contingence et calculs statistiques par item - Protocoles A et B

Item 1 : Contenance labiale

N.C. : Non Calculable

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	3	9	12
	T-	2	18	20
	Total	5	27	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	0	0	0
	T-	5	27	32
	Total	5	27	32

Se(A)	0,6
-------	-----

Se(B)	0
-------	---

Bornes inf / Borne sup	0,170585515	1,029414485
IC(SeA)95%	= [0,17 ; 1,02]	

Bornes inf / Borne sup		
IC(SeB)95%		N.C.

Spe(A)	0,666666667
--------	-------------

Spe(B)	1
--------	---

Bornes inf / Borne sup	0,488851856	0,844481478
IC(SpeA)95%	= [0,49 ; 0,84]	

Bornes inf / Borne sup		
IC(SpeB)95%		N.C.

J(A)	0,26666667
------	------------

J(B)	0
------	---

Item 2 : Mouvements mandibulaires

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	5	5	10
	T-	5	17	22
	Total	10	22	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	4	3	7
	T-	6	19	25
	Total	10	22	32

Se(A)	0,5
-------	-----

Se(B)	0,4
-------	-----

Bornes inf / Borne sup	0,190096789	0,809903211
IC(SeA)95%	= [0,19 ; 0,81]	

Bornes inf / Borne sup	0,096358106	0,703641894
IC(SeB)95%	= [0,10 ; 0,70]	

Spe(A)	0,772727273
--------	-------------

Spe(B)	0,863636364
--------	-------------

Bornes inf / Borne sup	0,597608981	0,947845565
IC(SpeA)95%	= [0,60 ; 0,95]	

Bornes inf / Borne sup	0,720232954	1,007039773
IC(SpeB)95%	= [0,72 ; 1,01]	

J(A)	0,27272727
------	------------

J(B)	0,26363636
------	------------

ANNEXE 7 (suite)

Item 3 : Rythme de succion-déglutition

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	14	4	18
	T-	4	10	14
	Total	18	14	32

Se(A)	0,777777778	
Bornes inf / Borne sup	0,585715831	0,969839725
IC(SeA)95%	= [0,59 ; 0,97]	

Spe(A)	0,714285714	
Bornes inf / Borne sup	0,477642523	0,950928906
IC(SpeA)95%	= [0,48 ; 0,95]	

J(A)	0,49206349
------	------------

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	9	2	11
	T-	9	12	21
	Total	18	14	32

Se(B)	0,5	
Bornes inf / Borne sup	0,269011785	0,730988215
IC(SeB)95%	= [0,27 ; 0,73]	

Spe(B)	0,857142857	
Bornes inf / Borne sup	0,673839829	1,040445885
IC(SpeB)95%	= [0,67 ; 1,04]	

J(B)	0,35714286
------	------------

Item 4 : Mouvements linguaux

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	6	12	18
	T-	1	13	14
	Total	7	25	32

Se(A)	0,857142857	
Bornes inf / Borne sup	0,597913229	1,116372485
IC(SeA)95%	= [0,60 ; 1,12]	

Spe(A)	0,52	
Bornes inf / Borne sup	0,324156863	0,715843137
IC(SpeA)95%	= [0,32 ; 0,72]	

J(A)	0,37714286
------	------------

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	2	0	2
	T-	5	25	30
	Total	7	25	32

Se(B)	0,285714286	
Bornes inf / Borne sup	-0,04894972	0,620378296
IC(SeB)95%	= [-0,05 ; 0,62]	

Spe(B)	1	
Bornes inf / Borne sup		
IC(SpeB)95%		N.C.

J(B)	0,28571429
------	------------

ANNEXE 7 (suite)

Item 5 : Initiation de la déglutition pharyngée

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	12	1	13
	T-	13	6	19
	Total	25	7	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	2	0	2
	T-	23	7	30
	Total	25	7	32

Se(A)	0,48
-------	------

Se(B)	0,08
-------	------

Bornes inf / Borne sup	0,284156863	0,675843137
IC(SeA)95%	= [0,28 ; 0,68]	

Bornes inf / Borne sup	-0,02634693	0,186346934
IC(SeB)95%	= [-0,03 ; 0,19]	

Spe(A)	0,857142857
--------	-------------

Spe(B)	1
--------	---

Bornes inf / Borne sup	0,597913229	1,116372485
IC(SpeA)95%	= [0,60 ; 1,12]	

Bornes inf / Borne sup		
IC(SpeB)95%		N.C.

J(A)	0,33714286
------	------------

J(B)	0,08
------	------

Item 6 : Propulsion pharyngée

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	7	4	11
	T-	8	13	21
	Total	15	17	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	2	0	2
	T-	13	17	30
	Total	15	17	32

Se(A)	0,466666667
-------	-------------

Se(B)	0,133333333
-------	-------------

Bornes inf / Borne sup	0,214194681	0,719138653
IC(SeA)95%	= [0,21 ; 0,72]	

Bornes inf / Borne sup	-0,03869733	0,305363994
IC(SeB)95%	= [-0,04 ; 0,31]	

Spe(A)	0,764705882
--------	-------------

Spe(B)	1
--------	---

Bornes inf / Borne sup	0,56306232	0,966349445
IC(SpeA)95%	= [0,56 ; 0,97]	

Bornes inf / Borne sup		
IC(SpeB)95%		N.C.

J(A)	0,23137255
------	------------

J(B)	0,13333333
------	------------

ANNEXE 7 (suite et fin)

Item 7 : Contenance vélopharyngée

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	0	1	1
	T-	5	26	31
	Total	5	27	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	0	0	0
	T-	5	27	32
	Total	5	27	32

Se(A)	0
Bornes inf / Borne sup	
IC(SeA)95%	N.C.

Se(B)	0
Bornes inf / Borne sup	
IC(SeB)95%	N.C.

Spe(A)	0,962962963
Bornes inf / Borne sup	0,891727361 1,034198565
IC(SpeA)95%	= [0,89 ; 1,03]

Spe(B)	1
Bornes inf / Borne sup	
IC(SpeB)95%	N.C.

J(A)	-0,037037
------	-----------

J(B)	0
------	---

Item 8 : Protection des voies respiratoires

Protocole A		VRDG		
		M+	M-	Total
Grille	T+	2	7	9
	T-	2	21	23
	Total	4	28	32

Protocole B		VRDG		
		M+	M-	Total
Grille	T+	0	1	1
	T-	4	27	31
	Total	4	28	32

Se(A)	0,5
Bornes inf / Borne sup	0,01 0,99
IC(SeA)95%	= [0,01 ; 0,99]

Se(B)	0
Bornes inf / Borne sup	
IC(SeB)95%	N.C.

Spe(A)	0,75
Bornes inf / Borne sup	0,589609851 0,910390149
IC(SpeA)95%	= [0,59 ; 0,91]

Spe(B)	0,964285714
Bornes inf / Borne sup	0,895547079 1,03302435
IC(SpeB)95%	= [0,90 ; 1,03]

J(A)	0,25
------	------

J(B)	-0,0357143
------	------------

Titre : Etude de validité de critère d'une grille d'évaluation du nourrisson lors de la prise alimentaire au biberon

Résumé

L'incidence des troubles de l'alimentation et de la déglutition dans la population pédiatrique a augmenté ces dernières années. Le besoin d'outils d'évaluation adaptés et standardisés se fait donc ressentir chez les professionnels de santé de la petite enfance. Une grille d'observation clinique de la succion et de la déglutition des nouveau-nés a ainsi été créée et en partie validée par les précédents travaux de Pauline Crusson (2013), Morgane Péron (2018) et Camille Poumarède (2019). L'objectif de la présente étude est ainsi de poursuivre la standardisation de cet outil en étudiant sa validité de critère. Ceci doit permettre aux orthophonistes et autres professionnels de la petite enfance de poser un diagnostic fiable et moins coûteux que la vidéoradioscopie de la déglutition, sur le plan financier ainsi qu'en terme d'irradiation. Pour cela, les 11 mêmes enfants sont évalués lors de la tétée au biberon au moyen de la grille testée, ainsi qu'au moyen d'une vidéoradioscopie, examen de référence. Les résultats sont comparés au niveau global et plus spécifique. La validité de critère est ainsi démontrée concernant les scores globaux, mais n'est pas significative concernant la comparaison item par item. Cette étude comporte néanmoins des limites, notamment concernant le choix de la population, du Gold Standard et des procédures de passation. Des adaptations sont donc envisageables afin d'améliorer les capacités diagnostiques de ce test (échantillonnage, mode de cotation, poursuite de la standardisation) et de le perfectionner en vue d'une utilisation à plus grande échelle dans le domaine des troubles de l'alimentation et de la déglutition des nourrissons.

Mots-clés : déglutition, succion, nourrissons, allaitement au biberon, grille d'observation, validité

Title : Criterion validity study of an assessment scale for infants during bottle-feeding

Abstract

The incidence of feeding and swallowing disorders in the pediatric population has increased in recent years. Adapted and standardized assessment tools are therefore needed by health early childhood professionals. A clinical observation scale for sucking and swallowing in newborns population has thus been created and partly validated by the previous works of Pauline Crusson (2013), Morgane Péron (2018) and Camille Poumarède (2019). The objective of the present study is to pursue the standardization of this tool by studying its criterion validity. This should allow speech therapists and other early childhood professionals to make a reliable and less demanding diagnosis than videofluoroscopy of swallowing, financially as well as in terms of radiations. For this purpose, the same 11 children were evaluated during bottle-feeding using both a tested grid, and the technics of videofluoroscopy, which remains the reference examination. The results are compared at a global and more specific level. The criterion validity is thus demonstrated concerning the global scores, but is not significant concerning the item by item comparison. This study nevertheless has limitations, particularly concerning the choice of the population, the Gold Standard and the test procedures. Adaptations are therefore possible in order to improve the diagnostic capacities of this test (sampling, scoring method, further standardization) and to improve it for a wider use in the field of feeding and swallowing disorders in infants.

Keywords : swallowing, sucking, infants, bottle-feeding, observation scale, criterion validity