

Mémoire de fin d'études



Institut Limousin de FOrmation aux MÉtiers de la Réadaptation Orthophonie

L'apport des modèles d'apprentissage dans les étalonnages
Poursuite de l'étalonnage du QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire)

Mémoire présenté et soutenu par
Chiara Treyt

En juin 2025

Mémoire dirigé par
Olivier Prot
Maître de conférences



Remerciements

Je tiens à exprimer mes sincères remerciements à toutes les personnes qui m'ont accompagnée tout au long de la réalisation de ce mémoire.

Tout d'abord, je souhaite remercier M. Olivier Prot, mon directeur de mémoire, pour son investissement et son aide dans ce travail. Il s'est toujours montré très disponible et a su me guider avec bienveillance. Son expertise et son soutien ont été essentiels à l'aboutissement de ce projet.

Je remercie l'ensemble de l'équipe pédagogique de l'ILFOMER et tout particulièrement Mme Audrey Pépin-Boutin pour ses conseils et sa disponibilité notamment durant cette dernière année d'étude qui a été très dense.

Je remercie chaleureusement Louise Barral, Léa Bardousse, Pauline Grueau, Hélène Dumonteil et Eléonore Gendreau pour leurs travaux précédant le mien et pour leur aide à la diffusion du questionnaire.

Merci à tous les professionnels de santé ayant diffusé le questionnaire, et à tous les parents ayant accepté d'y répondre. Je remercie spécifiquement Mme Justine Pradeau-Gratadour, orthophoniste m'ayant accueillie en stage, qui s'est particulièrement investie dans la recherche de participants à mon étude.

Je souhaite également remercier les membres du jury qui ont accepté de lire et d'évaluer ce mémoire.

Merci à Maëlle pour son soutien tout au long de ces études, autant dans le travail que dans tous les autres moments partagés.

Enfin et surtout, un grand merci à ma famille, notamment à ma maman et à mes frères et sœurs, pour leur soutien et leurs sourires à toute épreuve. Merci aussi à Alban, de m'avoir supportée et soutenue.

Droits d'auteurs

Cette création est mise à disposition selon le Contrat :

« **Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** »

disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



Charte anti-plagiat

La Direction Régionale de la Jeunesse, des Sports et de la Cohésion Sociale délivre sous l'autorité du Préfet de région les diplômes du travail social et des auxiliaires médicaux et sous l'autorité du Ministre chargé des sports les diplômes du champ du sport et de l'animation.

Elle est également garante de la qualité des enseignements délivrés dans les dispositifs de formation préparant à l'obtention de ces diplômes.

C'est dans le but de garantir la valeur des diplômes qu'elle délivre et la qualité des dispositifs de formation qu'elle évalue que les directives suivantes sont formulées à l'endroit des étudiants et stagiaires en formation.

Article 1 :

Tout étudiant et stagiaire s'engage à faire figurer et à signer sur chacun de ses travaux, deuxième de couverture, l'engagement suivant :

Je, soussigné Chiara Treyt

**atteste avoir pris connaissance de la charte anti plagiat élaborée par la DRDJSCS NA
– site de Limoges et de m'y être conformé.**

Et certifie que le mémoire/dossier présenté étant le fruit de mon travail personnel, il ne pourra être cité sans respect des principes de cette charte.

Fait à Pessac, Le jeudi 1er mai 2025

Suivi de la signature.



Article 2 :

« Le plagiat consiste à insérer dans tout travail, écrit ou oral, des formulations, phrases, passages, images, en les faisant passer pour siens. Le plagiat est réalisé de la part de l'auteur du travail (devenu le plagitaire) par l'omission de la référence correcte aux textes ou aux idées d'autrui et à leur source ».

Article 3 :

Tout étudiant, tout stagiaire s'engage à encadrer par des guillemets tout texte ou partie de texte emprunté(e) ; et à faire figurer explicitement dans l'ensemble de ses travaux les références des sources de cet emprunt. Ce référencement doit permettre au lecteur et correcteur de vérifier l'exactitude des informations rapportées par consultation des sources utilisées.

Article 4 :

Le plagitaire s'expose aux procédures disciplinaires prévues au règlement intérieur de l'établissement de formation. Celles-ci prévoient au moins sa non présentation ou son retrait de présentation aux épreuves certificatives du diplôme préparé.

En application du Code de l'éducation et du Code pénal, il s'expose également aux poursuites et peines pénales que la DRJSCS est en droit d'engager. Cette exposition vaut également pour tout complice du délit.

Vérification de l'anonymat

Mémoire Certificat de Capacité d'Orthophonie

Session de juin 2025

Attestation de vérification d'anonymat

Je soussignée(e) Chiara Treyt

Etudiant.e de 5ème année

Atteste avoir vérifié que les informations contenues dans mon mémoire respectent strictement l'anonymat des personnes et que les noms qui y apparaissent sont des pseudonymes (corps de texte et annexes).

Si besoin l'anonymat des lieux a été effectué en concertation avec mon Directeur de mémoire.

Fait à : Pessac

Le : jeudi 1er mai 2025

Signature de l'étudiant.e



Glossaire

ACP : Analyse en Composantes Principales

IMC : Indice de Masse Corporelle

QUINOA : Questionnaire d'investigation de l'oralité alimentaire

TAP : Trouble alimentaire pédiatrique

VPN : Valeur prédictive négative

VPP : Valeur prédictive positive

Table des matières

Introduction	12
Cadre théorique.....	13
1. L'histoire du QUINOA	13
1.1. L'oralité et l'oralité alimentaire	13
1.2. L'oralité alimentaire troublée : les troubles alimentaires pédiatriques (TAP)	13
1.3. Le dépistage et le diagnostic des troubles alimentaires pédiatriques	13
1.4. Le QUINOA	15
2. Les étalonnages	15
2.1. Définition et objectif d'un étalonnage	15
2.2. Classification binaire avec un score linéaire	15
2.3. Classification binaire par régression logistique	16
2.4. Analyse en composantes principales.....	16
2.5. Les métriques d'un outil de décision.....	17
2.5.1. La sensibilité	17
2.5.2. La spécificité	17
2.5.3. La précision	18
2.5.4. Les valeurs prédictives positives et négatives	18
2.6. La courbe ROC	19
2.7. La variance et les biais dans le cadre de l'apprentissage statistique	20
2.7.1. Le sous-apprentissage et le sur-apprentissage	20
2.7.2. La régularisation	21
2.8. Les différents types d'étalonnage	21
2.8.1. Quantilages et échelles normalisées	21
2.8.2. Généralités sur les modèles d'apprentissage	22
2.8.3. Etude d'un modèle d'apprentissage supervisé : le modèle de régression logistique	23
2.9. L'aspect psychométrique	24
Objectif de recherche	26
Protocole expérimental.....	27
1. Aspect éthique : formulaire de consentement	27
2. Population.....	27
3. Variables.....	28
4. Participants	28
5. Matériel.....	28
5.1. Version consensuelle du QUINOA.....	29
5.2. Questions ajoutées.....	29
5.3. Consignes de passation du QUINOA	29
6. Diffusion du QUINOA.....	30
6.1. Support d'administration	30
6.2. Recrutement.....	30
7. Méthode d'analyse des résultats.....	31
Présentation des résultats	32
1. Questionnaires inclus dans l'étude.....	32
2. Participants	32
3. Temps de passation.....	32

4. Résultats obtenus	33
5. Calcul du score	33
6. Distribution des scores obtenus	33
6.1. Statistiques globales.....	33
6.2. ACP.....	34
7. Détection d'un patient susceptible de présenter un TAP	35
7.1. Les méthodes de Gendreau (2024) : le seuil d'alerte.....	35
7.1.1. Le seuil d'alerte déterminé à partir du z-score.....	35
7.1.2. Le seuil d'alerte déterminé à partir des quantiles	36
7.2. Une autre méthode : la régression logistique.....	36
8. Les métriques des outils de décision.....	37
8.1. Le seuil d'alerte selon Gendreau (2024).....	38
8.2. La régression logistique.....	38
8.3. Visualisation des deux détecteurs	38
8.4. Courbes ROC.....	40
9. Régularisation l1	40
10. L'équilibre biais/variance du modèle de régression logistique	41
11. Les aspects psychométriques du QUINOA	42
11.1. La cohérence interne	42
11.2. La fiabilité test-retest	42
11.3. La fidélité interjuges.....	42
Discussion	43
1. Questions relatives à une prise en soin.....	43
2. Comparaison des méthodes	43
2.1. Synthèse des données par l'ACP	43
2.2. Discussion des méthodes de Gendreau (2024)	43
2.3. Discussion du modèle de régression logistique	44
2.4. Comparaison des deux méthodes	45
2.5. Les aspects psychométriques du QUINOA.....	45
3. Biais méthodologiques et limites de l'étude	45
4. Intérêts de l'étude	46
5. Perspectives du travail autour du QUINOA	47
5.1. Poursuite de l'étalonnage du QUINOA	47
5.2. Poursuite du recueil des mesures psychométriques du QUINOA	47
Conclusion	48
Références bibliographiques	49
Sitographie	53
Annexes	54

Table des illustrations

Figure 1 : Exemples de courbes ROC	19
Figure 2 : Répartition des participants en fonction de l'âge.....	32
Figure 3 : Temps de passation du questionnaire, en secondes	33
Figure 4 : Distribution des scores obtenus.....	34
Figure 5 : Premier graphique de l'ACP des scores obtenus aux deux phases d'étalonnage du QUINOA.....	34
Figure 6 : Deuxième graphique de l'ACP des scores obtenus aux deux phases d'étalonnage du QUINOA	35
Figure 7 : Graphique du point de vue du score de Gendreau (2024)	39
Figure 8 : Graphique du point de vue de la régression logistique	39
Figure 9 : Courbes ROC du modèle de Gendreau (2024) et du modèle de régression logistique	40
Figure 10 : Réponses à la question 1	85
Figure 11 : Réponses à la question 1bis.....	85
Figure 12 : Réponses à la question 1ter	85
Figure 13 : Réponses à la question 2	86
Figure 14 : Calcul de l'IMC selon les réponses aux questions 2, 3 et 4.....	86
Figure 15 : Réponses à la question 5	86
Figure 16 : Réponses à la question 6	87
Figure 17 : Réponses à la question 7	87
Figure 18 : Réponses à la question 8	87
Figure 19 : Réponses à la question 9	88
Figure 20 : Réponses à la question 10	88
Figure 21 : Réponses à la question 11	88
Figure 22 : Réponses à la question 12	89
Figure 23 : Réponses à la question 13	89
Figure 24 : Réponses à la question 14	90
Figure 25 : Réponses à la question 15	90
Figure 26 : Réponses à la question 16	90
Figure 27 : Réponses à la question 17	91
Figure 28 : Réponses à la question 18	91
Figure 29 : Réponses à la question 19	91
Figure 30 : Réponses à la question 20	92

Figure 31 : Réponses à la question 21	92
Figure 32 : Réponses à la question 22	92
Figure 33 : Réponses à la question 23	93
Figure 34 : Réponses à la question 24	93
Figure 35 : Réponses à la question 25	93
Figure 36 : Réponses à la question 26	94

Table des tableaux

Tableau 1 : Exemple d'un outil de décision.....	17
Tableau 2 : Comparaison des résultats entre la régression logistique et la méthode de Gendreau (2024) sur le jeu test	37
Tableau 3 : Résumé des résultats d'après la méthode de Gendreau (2024) et d'après le modèle de la régression logistique sur le jeu test	38
Tableau 4 : Répartition des points pour la question 5	71
Tableau 5 : Répartition des scores pour la question 6	71
Tableau 6 : Répartition des points pour la question 7	71
Tableau 7 : Répartition des points pour la question 8	71
Tableau 8 : Répartition des points pour la question 9	71
Tableau 9 : Répartition des points pour la question 10	72
Tableau 10 : Répartition des points pour la question 11	72
Tableau 11 : Répartition des points pour la question 12	72
Tableau 12 : Répartition des points pour la question 13	72
Tableau 13 : Répartition des points pour la question 14	72
Tableau 14 : Répartition des points pour la question 15	73
Tableau 15 : Répartition des points pour la question 16	73
Tableau 16 : Répartition des points pour la question 17	73
Tableau 17 : Répartition des points pour la question 18	73
Tableau 18 : Répartition des points pour la question 19	73
Tableau 19 : Répartition des points pour la question 20	73
Tableau 20 : Répartition des points pour la question 21	74
Tableau 21 : Répartition des points pour la question 22	74
Tableau 22 : Répartition des points pour la question 23	74
Tableau 23 : Répartition des points pour la question 24	74
Tableau 24 : Répartition des points pour la question 25	74
Tableau 25 : Répartition des points pour la question 26	75

Introduction

Un étalonnage consiste à faire passer un test à deux populations : une saine et une porteuse d'un trouble, et à tirer des conclusions concernant les différents comportements relevés (Piéron, 1951). Il existe différentes manières de réaliser un étalonnage : d'une façon plutôt classique avec le calcul d'un score seuil, à une manière plus moderne à l'aide d'un modèle d'apprentissage. L'étalonnage est une étape obligatoire pour qu'un outil de décision (test dépistage ou test diagnostic par exemple) puisse être considéré comme valide (Loubert et al., 2020).

Le QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire) est un test de dépistage des troubles alimentaires pédiatriques (TAP). Un TAP représente l'ensemble des difficultés en lien avec l'alimentation orale (Vidal, 2015). Le QUINOA est donc un questionnaire ayant pour objectif de déterminer si un enfant tout-venant, âgé de 1 an à 6 ans, est potentiellement porteur de ce trouble, ou non. Cet outil de décision est en création depuis 2018, dans le cadre de mémoires d'orthophonie. C'est Louise Barral qui a imaginé et amorcé ce projet. Depuis, il a été repris par 5 autres étudiantes du centre de formation de Limoges. En 2024, Eléonore Gendreau a réalisé la première partie de l'étalonnage du QUINOA, soit celle concernant les enfants qui ne présentaient pas de TAP. Nous avons fait le choix de poursuivre ce travail, qui permettra de combler le manque crucial d'un test de dépistage des TAP chez les enfants français âgés de 1 an à 6 ans. Il n'existe actuellement pas de tel test, alors que près d'un quart des nourrissons et des jeunes enfants présentent des difficultés alimentaires (Dubedout et al., 2016). Un dépistage précoce permettra une prise en soin précoce, qui aura pour objectif de prévenir les difficultés relevées et ainsi essayer de réduire l'impact du trouble sur le développement de l'enfant (Thibault, 2017).

Ainsi, notre travail aura pour objectif de réaliser un recueil de réponses du QUINOA, auprès d'enfants âgés de 1 an à 6 ans, porteurs d'un TAP. Pour ce faire, nous contacterons via divers réseaux, des professionnels de santé susceptibles de prendre en soin des enfants porteurs d'un TAP. Puis nous traiterons ces données grâce à un modèle d'apprentissage.

Ce projet contribue donc à la création du QUINOA, un outil de dépistage des TAP chez les enfants de 1 an à 6 ans, à destination des professionnels de santé.

Pour mener à bien ce projet, nous présenterons d'un point de vue théorique les éléments clés de notre sujet : l'histoire du QUINOA puis les étalonnages. Par la suite, nous détaillerons notre objectif de recherche puis le protocole expérimental que nous allons mettre en place pour aboutir à la réalisation de notre objectif de recherche. Puis nous présenterons les résultats de notre étude, et nous les commenterons. Pour finir, nous conclurons et nous présenterons les sources issues de la littérature qui nous ont permis de mener à bien ce projet.

Cadre théorique

1. L'histoire du QUINOA

1.1. L'oralité et l'oralité alimentaire

L'oralité renvoie à toutes les fonctions en lien avec la sphère buccale, telle que la communication et l'alimentation (Abadie, 2004). L'oralité alimentaire désigne le développement alimentaire de chaque individu et l'oralité verbale renvoie aux actes de communication. L'oralité alimentaire est scindée en deux phases : l'oralité primaire et l'oralité secondaire. L'oralité primaire est régie par le tronc cérébral et est un acte automatique qui comprend la succion-déglutition et l'oralité secondaire se développe après quelques mois de vie, notamment avec le passage à la cuillère et le développement de la mastication qui sont mis en place avec la diversification alimentaire (Thibault, 2015). L'oralité alimentaire constitue le lien entre le bébé et le monde qui l'entoure (Rofidal, 2022). Cette oralité commence son développement in-utero très tôt : dès le 4ème mois de grossesse, le futur bébé commence à déglutir et dès la 15 semaines la succion automatique se met en place (Quentin, 2015). Cette coordination automatique in-utero va permettre au nourrisson d'entreprendre des expériences olfactives, gustatives et tactiles, ce qui va développer son oralité alimentaire (Bellis et al., 2009).

1.2. L'oralité alimentaire troublée : les troubles alimentaires pédiatriques (TAP)

20 à 25% des nourrissons et des jeunes enfants présentent des difficultés alimentaires (Dubedout et al., 2016). D'après Goulet et al., en 2012, 3 à 10% des enfants présentent réellement un trouble alimentaire pédiatrique. Un trouble alimentaire pédiatrique correspond à la perturbation de l'apport oral de nutriments d'une durée d'au moins 2 semaines et associée à un dysfonctionnement médical et/ou un dysfonctionnement nutritionnel et/ou un dysfonctionnement de la capacité d'alimentation et/ou un dysfonctionnement psychosocial (Goday et al., 2019). La prise en soin orthophonique est nécessaire dans le cadre d'un trouble alimentaire pédiatrique afin de travailler sur les compétences fonctionnelles oromotrices de l'enfant, mais également sur les aspects comportementaux et sensoriels (Mottrie et al., 2022). Abadie précise en 2004, que l'oralité primaire peut être stimulée mais ne pourra pas être rééduquée, c'est pour cela qu'il faut que la sphère orofaciale soit protégée au mieux des éléments négatifs qui pourraient influer sur son bon fonctionnement.

L'orthophoniste, qui est un professionnel de santé évaluant et prenant en soin des troubles du langage, de la communication, des fonctions oro-myo-faciales, de la phonation, de la cognition mathématique, est également concerné par le développement de l'oralité. Il peut, en cas de besoin, effectuer un bilan et une rééducation d'un TAP. A la suite du bilan, l'orthophoniste détermine si le patient est porteur d'un TAP ou non.

1.3. Le dépistage et le diagnostic des troubles alimentaires pédiatriques

Certains éléments dans le développement de l'enfant sont considérés comme des signes d'alerte d'un TAP. Grevesse et al., en 2020, les répertorient en différentes catégories :

- Les signes comportementaux tels qu'une cambrure ou un raidissement du corps pendant l'alimentation, une irritabilité ou un manque de vigilance lors des repas, un refus de la nourriture solide ou liquide, des troubles du comportement au repas, des conflits lors des repas, de la néophobie (peur ou refus d'expérimenter des aliments

nouveaux ou inconnus), des interruptions des repas, un manque d'autonomie alimentaire, l'utilisation d'ustensiles inappropriés à l'âge (par exemple boire au biberon à 4 ans) ;

- Les signes fonctionnels tels qu'un retard de développement de la motricité orale, une incapacité à gérer en bouche différentes textures de nourriture, des difficultés de succion et/ou de mastication, une toux pendant les repas, un bavage excessif, de la nourriture solide ou liquide sortant de la bouche ou du nez, des difficultés à coordonner la respiration en mangeant et en buvant, un étouffement lors des repas, des infections respiratoires récidivantes ;
- Les signes sensoriels tels que toutes les formes de sélectivité alimentaire (par exemple aux textures, aux couleurs), un enfant qui mange mieux les aliments préférés (ce qui peut être qualifié à tort de caprice), la présence de désordres sensoriels non alimentaires (par exemple des enfants qui ne supportent pas de toucher certaines textures) ;
- Les signes mixtes tels que des régurgitations et vomissements, un allongement de la durée des repas (au-delà de 30 minutes), des grimaces, la présence d'un réflexe nauséeux, le refus de poursuivre le repas, une détresse en cas de forçage.

Quelques outils ont été créés permettant de dépister un TAP, tel que ORALQUEST qui est un questionnaire élaboré par le service de pédiatrie de l'Hôpital Necker-Enfants malades de Paris (Hôpital Necker-Enfants Malades, s.d.). Ce questionnaire a pour vocation de déceler les TAP chez des enfants porteurs d'une pathologie congénitale. La publication de ce questionnaire était prévue pour 2020, n'est cependant pas encore validé ni publié. Ce n'est donc pas un outil sur lequel les professionnels de santé peuvent s'appuyer.

Le « grille parentale exploratrice pour des enfants âgés de 24 à 36 mois » est un autre outil de dépistage des TAP, pour une population âgée de 24 à 36 mois. Cette grille, qui a été élaborée dans le cadre d'un mémoire d'orthophonie, a été étalonnée sur 15 individus (Gelas et Monier-Guillaumin, 2016). Cette population est trop faible pour pouvoir considérer l'outil comme fiable.

L'échelle d'alimentation de l'Hôpital de Montréal pour enfants, nommée The Montreal Children's Hospital Feeding Scale est un outil de dépistage des TAP chez les enfants âgés de 6 mois à 83 mois (Ramsay et al., 2011). Ce test est disponible en anglais et en français. Il a été étalonné sur une population canadienne, ce qui pourrait engendrer des biais pour son utilisation en France. Ainsi, il pourrait être intéressant de réaliser un nouvel étalonnage de cet outil, sur une population française.

Le diagnostic de TAP, quant à lui, repose uniquement sur des observations cliniques, qui ont lieu sur les temps des repas. Les professionnels de santé accompagnant la famille ne sont généralement pas présents sur ces temps-là, ce qui explique que les TAP ne soient pas bien dépistés ni diagnostiqués. En dehors de ces recommandations d'observations répertoriées par Grevesse et al. en 2020, il n'existe pas d'autres recommandations ou tests de dépistage des TAP fiables à ce jour.

Plus le diagnostic de TAP sera posé tardivement, plus la rééducation de ce trouble sera retardée, ce qui entraînerait une chronicisation du trouble ainsi que des conséquences sur le développement de l'enfant (Lecoufle, 2020). Cela montre qu'il est primordial de diagnostiquer ce trouble de manière précoce, et cela peut être favorisé par un test de dépistage fiable à destination des professionnels de santé.

1.4. Le QUINOA

D'après les chiffres relevés précédemment, un TAP est donc considéré comme un trouble fréquent. Après avoir constaté qu'il n'existe actuellement pas de test de dépistage des TAP fiable utilisable en France, on peut penser qu'il est nécessaire d'en créer un afin d'orienter le plus précocement et au mieux les enfants porteurs de ce trouble.

En partant de ce constat, Louise Barral, actuellement orthophoniste à Limoges, avait pour projet de créer, dans le cadre de son mémoire de fin d'études, un questionnaire de dépistage des troubles alimentaires pédiatriques à destination des pédiatres, pour les enfants de 0 à 6 ans. Elle avait commencé ce travail en 2018. Depuis, il a été repris par Léa Bardousse en 2019, également par Pauline Grueau en 2020, puis par Hélène Demonteil en 2022 et enfin par Eléonore Gendreau en 2024. Le QUINOA, QUEStionnaire d'INvestigation de l'Oralité Alimentaire, est donc en cours de réalisation depuis 2018. Après maints ajustements, le QUINOA est désormais prêt à être étalonné chez les enfants de 1 an à 6 ans. Aujourd'hui, la première partie de l'étalonnage sur des patients sains a été effectuée. A long terme, le QUINOA a donc pour objectif de déterminer rapidement et avec fiabilité si un enfant tout-venant résidant en France est potentiellement porteur d'un TAP.

2. Les étalonnages

2.1. Définition et objectif d'un étalonnage

L'étalonnage, c'est « un barème utilisé pour le classement d'une valeur individuelle par rapport à l'ensemble des valeurs caractéristiques d'une population » (Piéron, 1951).

L'étalonnage va donc permettre de déterminer, pour une population donnée, où se situent les résultats obtenus par rapport à une norme, qui sera également définie par l'étalonnage. On s'intéresse donc à la classification des individus dans une catégorie.

Dans le cas du QUINOA, nous souhaitons classer les profils des patients selon un mode binaire, qui va correspondre à « porteur d'un TAP » ou « non-porteur d'un TAP ». Pour déterminer cette classification binaire, nous allons étudier deux méthodes distinctes.

2.2. Classification binaire avec un score linéaire

Il est possible de déterminer la classification binaire d'un étalonnage grâce à un score dit « linéaire ». Ce score peut être déterminé en calculant la somme des réponses obtenues, s'il s'agit de réponses quantitatives.

De manière générale, un score linéaire se calcule de la manière suivante :

$$S = \langle W, X \rangle + b$$

Où $\langle \cdot, \cdot \rangle$ désignent le produit scalaire usuel

W désigne le vecteur des poids associés à chaque question

X est le vecteur contenant les réponses à chacune des questions

b est un nombre appelé le biais.

Comme nous venons de le voir, chaque question a un poids (ou coefficient). L'auteur du test peut choisir d'attribuer des poids différents à chaque question, afin de mettre en avant les

éléments saillants de son étude ou au contraire, que toutes les questions aient le même poids, comme l'a fait Gendreau (2024) dans son étude. Ces poids vont avoir un impact sur le score total. Le fait que ces poids soient déterminés par un individu génère un biais, car ils sont déterminés de manière subjective. C'est pour cela que, généralement, les questions d'un étalonnage ne sont pas pondérées, afin que toutes les questions aient le même poids dans le calcul du score total dans le but d'éviter la subjectivité.

Le travail de Gendreau (2024) à propos du QUINOA était de transformer les réponses qualitatives en réponses quantitatives, pour établir un score linéaire. Cela lui a permis d'obtenir un score seuil, à partir duquel le profil étudié est placé dans l'une des catégories.

2.3. Classification binaire par régression logistique

Le modèle de régression logistique, très couramment utilisé en pratique (Hastie et al., 2009), permet de réaliser une classification binaire, comme dans le cas de notre étude. Il s'agit d'un modèle d'apprentissage supervisé, qui sera détaillé par la suite.

La régression logistique utilise, pour classer les données, un score tel que défini dans la section précédente, qui est déterminé par les poids associés à chaque question. Ces poids ne sont pas choisis de manière subjective par une personne, mais sont déterminés lors du processus d'apprentissage et représente les meilleurs choix possibles pour que le modèle explique au mieux le jeu de données d'apprentissage.

2.4. Analyse en composantes principales

Actuellement, on assiste à une quantité de données grossissantes (Comité consultatif national d'éthique, 2019), qui sont parfois difficilement exploitables. Les chercheurs rencontrent des difficultés à les traiter au vu du nombre de données. Le risque d'avoir des données imposantes, est de ne pas percevoir l'information pertinente.

Harold Hotelling, en 1933, a inventé l'Analyse en Composantes Principales (ACP) (Farebrother, 2022). Il s'agit de l'un des moyens existants pour résumer des données et les analyser plus facilement.

L'idée principale est de réduire les dimensions pour produire un graphique, habituellement en 2 dimensions, ou plus. L'objectif d'une ACP est donc de résumer l'information en créant un graphique synthétique et de permettre au lecteur de visualiser les corrélations possibles entre les variables (Saporta et Niang, 2003).

On suppose qu'il y a n mesures en entrée, ce qui correspond à un tableau à n dimensions. Cependant, si n est plus grand que 2 ou 3, il est impossible de faire un graphique ou un tableau lisible, et par suite de visualiser l'information pertinente. Pour faire une ACP, on détermine les composantes principales, c'est-à-dire les directions, dans l'espace à n dimensions des données, où la variance du diagramme de dispersion est maximisée. Cela produit donc un diagramme de dispersion suivant les directions principales et on peut également projeter sur ce graphique les axes correspondant aux entrées du tableau originel. Le centre du graphique correspond à la moyenne des données.

Pour récapituler, les avantages de l'ACP sont qu'elle permet une représentation plus simple des données et la détection des groupes similaires ou valeurs aberrantes (Saporta et Niang, 2003).

Dans la conception du QUINOA, nous serons amenés à utiliser des ACP, car le QUINOA comporte 22 questions, soit 22 dimensions. Ainsi, l'ACP nous permettra de visualiser l'ensemble des données de manière synthétique, tout en mettant en avant les corrélations entre les variables.

2.5. Les métriques d'un outil de décision

Pour imager ces différents concepts, nous allons prendre l'exemple d'un test de dépistage qui a été réalisé sur un échantillon de 100 personnes. Parmi ces 100 individus, nous savions que 35 étaient porteurs d'une maladie, et 65 n'étaient pas porteurs de cette maladie. Mais le test de dépistage a conclu que 50 personnes étaient porteuses de la maladie, et les autres n'étaient pas porteurs. Nous pouvons récapituler ces informations dans le tableau suivant :

Résultats du test	Porteurs de la maladie (+)	Non-porteurs de la maladie (-)	Total
Positif (+)	+ / + (Vrai positif) : 35	- / + (Faux positif) : 15	+ : 50
Négatif (-)	+ / - (Faux négatif) : 0	- / - (Vrai négatif) : 50	- : 50
Total	+ : 35	- : 65	100

Tableau 1 : Exemple d'un outil de décision

2.5.1. La sensibilité

Dans le cas d'un outil de décision, également appelé outil de détection, la sensibilité est la proportion de personnes qui sont détectées porteuses d'une pathologie et qui le sont réellement (Fawcett, 2006).

Plus le score de la sensibilité est proche de 1, plus cela signifie que l'outil de décision est sensible.

La sensibilité se calcule grâce à l'équation qui suit :

$$\text{Sensibilité (Se)} = \frac{\text{+} \text{+} (\text{Vrai positif})}{\text{+} (\text{Total des personnes porteuses de la maladie})}$$

Dans notre exemple, la sensibilité est donc égale à :

$$\text{Sensibilité (Se)} = \frac{35}{35}$$

$$= 1$$

La sensibilité est égale à 1 dans l'exemple donné, ce qui signifie que ce test est très sensible : il repère tous les patients qui sont porteurs d'une maladie.

2.5.2. La spécificité

La spécificité est la proportion de la population qui est considérée comme non porteuse de la pathologie d'après le test, et qui, en réalité, n'est effectivement pas porteuse de cette pathologie (Hastie et al., 2009).

Tout comme la sensibilité, plus le score de la spécificité est proche de 1, plus cela signifie que l'outil de décision est spécifique.

La spécificité se calcule grâce à l'équation qui suit :

$$\text{Spécificité (Sp)} = \frac{\text{Vrai négatif}}{\text{Total des personnes non-porteuses de la maladie}}$$

Dans notre exemple, la spécificité est donc égale à :

$$\begin{aligned}\text{Spécificité (Sp)} &= 50 / 65 \\ &= 0,78.\end{aligned}$$

Ce résultat signifie que le test a une bonne spécificité.

L'antispécificité désigne l'inverse de la spécificité, soit $1 - \text{Sp}$.

Dans notre exemple, l'antispécificité est égale à $1 - 0,78$, ce qui équivaut à 0,22.

2.5.3. La précision

La précision renvoie à la proportion des cas où le processus de décision ne donne pas une réponse erronée. Plus elle est élevée, plus le test est précis (Fawcett, 2006).

La précision se calcule grâce à l'équation suivante :

$$\text{Précision} = \frac{\text{Vrai positif} + \text{Vrai négatif}}{\text{Total des personnes porteuses de la maladie} + \text{Total des personnes non-porteuses de la maladie}}$$

Dans notre exemple, la précision est donc égale à :

$$\begin{aligned}\text{Précision} &= (35 + 50) / (35 + 65) \\ &= 0,85\end{aligned}$$

Au vu de ce résultat, nous pouvons dire que le test est assez précis.

2.5.4. Les valeurs prédictives positives et négatives

La valeur prédictive positive (notée VPP), est la probabilité, sachant que le test est positif, qu'une personne soit réellement porteuse de la maladie (Fawcett, 2006).

La valeur prédictive négative (notée VPN), représente la probabilité qu'une personne ne soit pas porteuse de la maladie, dans le cas où le test est négatif (Fawcett, 2006).

La valeur prédictive positive se calcule comme suit :

$$\text{VPP} = \frac{\text{Vrai positif}}{\text{Total des personnes ayant été détectées positives au test}}$$

Dans notre exemple, la VPP est donc égale à :

$$\begin{aligned}&= 35 / 50 \\ &= 0,70\end{aligned}$$

Ainsi, la VPP est égale à 0,70, ce qui signifie que si le test est positif, il y a 70% de probabilité pour que la personne soit porteuse de la maladie.

La valeur prédictive négative se calcule grâce à l'équation :

$$\text{VPN} = \frac{\text{Vrai négatif}}{\text{Total des personnes ayant été détectées négatives au test}}$$

Dans notre exemple, la VPN est donc égale à :

$$\begin{aligned} &= 50 / 50 \\ &= 1 \end{aligned}$$

La VPN est égale à 1, ce qui signifie que dans notre exemple, lorsque le test est négatif alors la probabilité que la personne ne soit pas porteuse de la maladie est de 100%.

Ces valeurs permettent donc de mettre en lumière les points positifs et négatifs d'un test. C'est un équilibre entre toutes ces valeurs qui détermine la pertinence d'un outil de décision (Fawcett, 2006).

2.6. La courbe ROC

La courbe ROC est une courbe qui établit le tracé de la sensibilité (en ordonnée) en fonction de l'antispécificité (en abscisse).

C'est un outil qui permet de comparer la qualité de deux outils de décision binaires. Des deux courbes représentant des outils de décision, celle positionnée au-dessus désigne l'outil le plus performant.

Voici un ensemble de courbes ROC qui va nous permettre d'expliciter le rapport entre les différentes courbes :

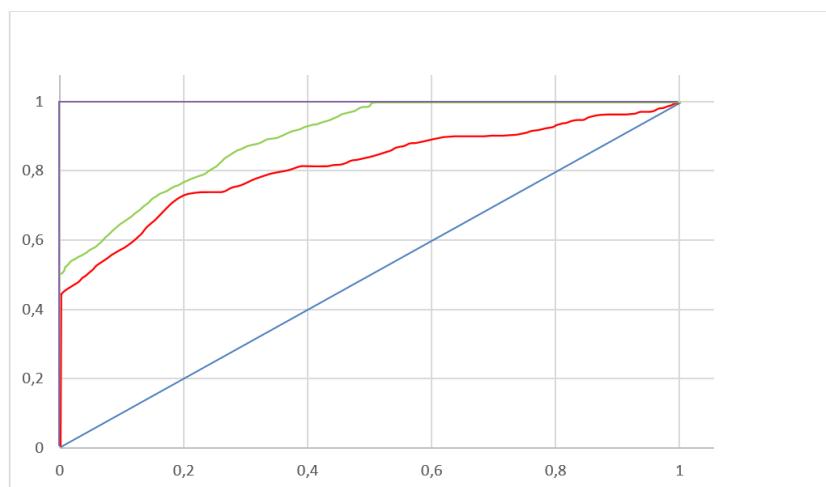


Figure 1 : Exemples de courbes ROC

Dans ce graphique, la courbe bleue correspond à l'outil qui décide au hasard. Les courbes rouge, verte et violette représentent des outils de détection. Nous pouvons voir que l'outil représenté par la courbe rouge est le moins fiable des trois, mais reste cependant plus fiable que le hasard. Celui représenté par la courbe violette est le meilleur outil des trois, et il désigne la courbe « parfaite », celle d'un outil le plus optimal possible.

2.7. La variance et les biais dans le cadre de l'apprentissage statistique

De manière générale, en apprentissage statistique on désigne par le biais, la précision du modèle pour effectuer correctement la tâche qui lui est demandée. Il ne faut pas confondre ce biais avec celui qui permet de calculer le score dans la formule vue précédemment, il s'agit de deux éléments distincts.

Pour un modèle linéaire, comme c'est le cas de la régression logistique par exemple, la variance correspond à la variance des poids utilisés dans la formule du score (notée W).

On peut montrer théoriquement (Hastie et al., 2009) que le biais et la variance sont toujours liés : c'est ce que l'on appelle *l'équilibre biais/variance*. En effet, si le biais diminue, la variance augmente. On peut penser qu'un modèle avec un très faible biais serait meilleur, mais cela signifierait que le modèle en question aurait des poids avec une très grande variance, donc des poids très fluctuants et qui seront finalement inutilisables.

Il convient donc de regarder l'équilibre biais/variance d'un modèle d'apprentissage pour décider si ce modèle est valide, en particulier vis-à-vis du sur-apprentissage comme nous le verrons dans la section suivante.

Dans le cadre du QUINOA, le biais désigne le fait que la régression logistique décide correctement si un enfant est porteur d'un TAP, ou pas.

2.7.1. Le sous-apprentissage et le sur-apprentissage

Il est possible qu'un modèle d'apprentissage sous-apprenne ou sur-apprenne. Le sous-apprentissage constitue le cas où un modèle est incapable d'avoir un biais faible. Autrement dit, et pour le cas d'un travail de classification, cela arrive lorsque le modèle est incapable de détecter correctement la bonne classe à partir des réponses aux questions.

Le sur-apprentissage, c'est lorsque le modèle a un biais très faible sur le jeu de données d'apprentissage mais que ce biais est beaucoup plus grand sur le jeu de données de test. Ainsi, le biais est très faible mais la variance est tellement grande que le modèle est complètement inefficace avec de nouvelles données (par exemple celle du jeu de test). Quand cela arrive de manière aussi extrême on dit que le modèle a appris par cœur (et c'est donc un modèle peu performant car il est incapable de généraliser ce qu'il a appris). C'est une des raisons pour laquelle on utilise un jeu de données test : la différence entre les métriques de performance sur le jeu test et le jeu d'entraînement sont une indication du sur- apprentissage.

L'étude de l'équilibre biais/variance est un moyen de détecter les cas de sur-apprentissage. Si on remarque qu'un modèle est en sur-apprentissage, cela signifie que sa variance est trop grande, et donc les poids de la formule du score ont une trop grande variabilité. Dans un tel cas, il est nécessaire de stabiliser le modèle. Pour ce faire, nous pouvons utiliser la régularisation.

Concernant le modèle de régression logistique, il est généralement peu sensible au sur-apprentissage.

2.7.2. La régularisation

Lorsqu'il y a sur-apprentissage d'un modèle, il est nécessaire de stabiliser le modèle. Il existe de très nombreuses techniques pour cela. Le sur-apprentissage vient en particulier du fait que le modèle est trop riche pour le problème considéré, ce qui conduit à une variance trop élevée. Une des techniques les plus utilisées est celle de la régularisation que nous allons présenter ici. Cette technique est opérante dans le cas où le modèle est en sur-apprentissage, mais elle permet également de sélectionner un modèle plus petit.

Pour éviter le sur-apprentissage d'un modèle, la régularisation utilise le terme de « pénalisation » pour stabiliser la variance des poids. Un coefficient de régularisation est associé à ce processus. Plus le coefficient sera grand, plus les poids seront stables mais plus le biais sera grand. A l'inverse, plus le coefficient sera petit, plus la variance sera grande mais le biais sera faible. Il convient donc de trouver un équilibre entre le biais et la variance à l'aide de ce coefficient.

Dans le cas particulier de la régularisation l_1 (aussi appelé Lasso pour un problème de régression) (Tibshirani, 1996), ce coefficient a également la propriété d'amener les poids à 0 dans le cas où la question n'est pas jugée importante par le modèle. Cela permet donc d'isoler les questions les plus importantes pour la prise de décision et d'avoir la possibilité de supprimer celles qui semblent peu pertinentes.

Dans le cas du QUINOA, dans lequel nous allons utiliser un modèle de régression logistique pour décider d'un potentiel TAP chez un enfant, il est connu que le risque de sur-apprentissage est faible. Nous n'utiliserons donc pas la régression pour stabiliser le modèle, mais plutôt pour mettre en lumière les questions les plus importantes.

2.8. Les différents types d'étalonnage

2.8.1. Quantilages et échelles normalisées

Nous allons désormais nous pencher sur le quantilage et les échelles normalisées qui sont deux manières utilisées classiquement pour étalonner les résultats.

Il existe deux modèles principaux en statistiques : le modèle paramétrique et le modèle non-paramétrique. Dans le modèle paramétrique, nous faisons des hypothèses sur la loi de probabilité inconnue qui est à l'origine des données. Dans ce cas, pour avoir connaissance des données, il suffit d'identifier les paramètres inconnus de cette loi de probabilité. Dans le modèle non-paramétrique, nous ne faisons aucune hypothèse sur la loi de probabilité. Ce modèle trouve un plus grand nombre d'applications pratiques, mais il se révèle moins précis.

Dans le cas d'un modèle non-paramétrique, nous utilisons souvent un quantilage afin d'identifier la loi de probabilité inconnue. Dans un quantilage, les réponses recueillies sont ordonnées par ordre croissant puis réparties en plusieurs classes. Ce type d'étalonnage, qui est davantage utilisé dans le cadre d'un petit échantillon récolté, peut être présenté sous forme de tableau ou de graphique. Les classes seront déterminées de telle manière qu'il y ait environ le même pourcentage de résultats dans chaque classe. Nous pourrons également faire apparaître des valeurs-seuils pour distinguer les cas pathologiques. Selon le nombre de classes définies, les quantilages portent différents noms : le décilage pour un étalonnage réparti en 10 classes, le quartilage pour un étalonnage en 4 classes, un centilage pour un étalonnage en 100 classes distinctes. Pour un décilage, on regroupe les résultats des participants en 10 classes distinctes, en faisant en sorte que chaque classe intègre environ

10% des résultats. Dans le cas d'un quartilage, les résultats sont répartis en 4 classes distinctes. Il faut donc que chaque classe représente, approximativement, 25% des résultats de la population observée. Enfin, pour le centilage, il s'agit de répartir les résultats obtenus en 100 classes distinctes. Le choix du nombre de classes dans le quantilage dépend de chaque étude et des résultats obtenus.

Dans le cas d'un modèle paramétrique, nous supposons généralement que nous cherchons une loi normale, ce qui explique que nous allons utiliser des échelles normalisées. Une loi normale suit la distribution théorique de Laplace-Gauss, qui est caractérisée par une forme de cloche : c'est une distribution continue et symétrique autour de sa moyenne m et telle que sa dispersion est caractérisée par son écart-type σ . Dans ce cas, nous avons en général besoin d'un grand échantillon pour mesurer correctement les paramètres m et σ de cette loi normale. Il convient de vérifier qu'il s'agisse bien d'une loi normale grâce au test de Shapiro-Wilk. Une fois que la moyenne et l'écart-type sont déterminés, il s'agit de déterminer les z-scores. Le z-score permet d'exprimer l'écart à la norme. Il s'exprime en écart-type. Le z-score se calcule par la formule suivante :

$$z\text{-score} = \frac{x - M}{\sigma}$$

Où x est le score brut

M est la moyenne

σ est l'écart-type

Dans le cas d'une loi normale, le z-score est utilisé pour différencier les personnes « saines » des personnes « pathologiques » : s'il est supérieur à 2 écarts-types ou inférieur à -2 écarts-types, alors la personne est considérée comme en dehors de la norme, ce qui la catégorise comme « pathologique », pour un alpha égal à 5%.

2.8.2. Généralités sur les modèles d'apprentissage

L'apprentissage statistique ou « machine learning » consiste à utiliser des données pour qu'une machine puisse entraîner un modèle d'apprentissage.

Il existe trois grands types d'apprentissage qui sont l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement.

L'apprentissage supervisé a pour objectif de prédire une valeur : c'est une tâche de *régression* ou de déterminer si les résultats obtenus démontrent une appartenance à un groupe : c'est une tâche de *classification*. Dans ce cas, on fournit des entrées et des sorties par le biais de jeu de données d'entraînement, et le but est d'entraîner le modèle afin de créer des liens entre ces entrées et ces sorties. Pour pouvoir mesurer les progrès dans l'apprentissage, le modèle dispose de différentes mesures pour évaluer ses performances sur ce jeu de données d'entraînement. L'apprentissage est supervisé car un superviseur donne les bonnes réponses, et le modèle doit apprendre de ces réponses pour pouvoir répondre correctement sur de nouvelles données (Hastie et al., 2009).

L'apprentissage non supervisé, quant à lui, se base sur l'étude de profils globaux. Ainsi, dans ce cas, le modèle d'apprentissage va permettre de rapprocher des individus en fonction de leur ressemblance globale (Hastie et al., 2009). Ici, il n'y a donc pas de superviseur pour

donner les bonnes réponses, ainsi c'est au modèle lui-même de tenter de mettre en lien certains profils grâce aux données recueillies.

Enfin, l'apprentissage par renforcement consiste en un travail par essai-erreur. En cas de succès, l'environnement fournit une récompense à la machine, ce qui la pousse à continuer dans cette voie et cela lui permet donc d'améliorer sa stratégie (Sutton, 2018). L'environnement est donc interactif, c'est-à-dire qu'il est modifié sous l'action de la machine.

Lorsqu'un modèle d'apprentissage est utilisé, on va scinder les résultats récoltés en 2 parties : la partie « entraînement », grâce à laquelle l'apprentissage du modèle est réalisé, et la partie « test » qui servira à valider la qualité de l'apprentissage. Cela va permettre en particulier d'éviter un sur-apprentissage ou un apprentissage « par cœur » de la machine, et donc de s'assurer de la fiabilité des résultats. En pratique, il est d'usage d'utiliser entre 70% et 80% de nos données pour l'entraînement et le reste pour la partie « test ». Il est essentiel de pouvoir mesurer les qualités et les défauts des modèles d'apprentissage, car tous les détecteurs ont des défauts qui, s'ils ne sont pas repérés, engendrent des erreurs et des inexactitudes dans le test en question.

Dans le cas du QUINOA, il semble pertinent d'utiliser un modèle d'apprentissage pour la classification afin de minimiser le pourcentage d'erreur. Il faudra donc reprendre les données recueillies par Gendreau (2024) concernant les résultats de la population qui n'est pas porteuse d'un TAP ainsi que collecter des données à propos de la population porteuse d'un TAP. Pour le QUINOA, nous pourrons donc utiliser l'apprentissage supervisé car nous serons en possession des données, la question restera donc seulement de déterminer si, au vu des résultats collectés, on suspecte un TAP ou non.

L'apprentissage supervisé rassemble différents modèles selon le problème posé. Dans notre étude, l'objectif est de prédire si une personne se situe dans la catégorie « porteuse d'un TAP » ou dans la catégorie « non porteuse d'un TAP » selon des variables indépendantes, ce qui sera possible grâce au modèle de régression logistique.

2.8.3. Etude d'un modèle d'apprentissage supervisé : le modèle de régression logistique

Le modèle de régression logistique est un modèle de classification, qui a pour objectif de prédire au mieux une variable dépendante à partir d'une ou plusieurs autres variables, celle(s)-ci indépendante(s). Ce modèle permet de vérifier la probabilité d'appartenance à une classe, comme la classe « sain » ou « malade », ou la probabilité qu'un phénomène se produise. La variable est dépendante de certains facteurs, qui peuvent être l'âge, la taille, la consommation de tabac, la prise alimentaire, etc.

Le modèle de régression logistique se base sur le lien suivant :

$$\log\left(\frac{P(C=1|X=x)}{P(C=0|X=x)}\right) = S = \langle W, X \rangle + b$$

Où P est la probabilité

C est la classe

S est le score

$\langle \cdot \rangle$ désignent le produit scalaire usuel

W désigne le vecteur des poids associé à chaque question

X est le vecteur contenant les réponses à chacune des questions

b est un nombre appelé le biais.

Dans le cas du QUINOA, nous sommes face à un problème de classification binaire. La classe, C, est donc égale à 1 si l'enfant est porteur d'un TAP, et à 0 dans le cas contraire.

2.9. L'aspect psychométrique

Dans un étalonnage, le versant psychométrique doit être abordé afin de s'assurer de la fiabilité du test ou du questionnaire, et donc savoir avec quel degré de précision il mesurera les variables. Il existe différents types de fiabilité : la cohérence interne, la fiabilité test-retest et la fidélité interjugés (Streiner et Norman, 1995).

La cohérence interne c'est le niveau de corrélation entre les différents items du test. La cohérence interne se base sur le principe que toutes les questions d'un test sont des indicateurs d'un même concept et donc qu'il est censé exister une relation significative entre ces dernières. La cohérence interne se mesure grâce au coefficient alpha de Cronbach.

L'alpha de Cronbach mesure la corrélation entre les questions, pour nous donner la précision de fiabilité du test. Il se situe entre 0 et 1. Plus la corrélation entre les items est élevée, plus la cohérence interne est élevée. D'après Streiner et Norman en 1995, pour qu'un test soit fiable, il faut que le coefficient alpha de Cronbach soit compris entre 0,70 et 0,90. Au-delà de 0,90, cela signifie que certaines questions sont redondantes, ce qui n'est pas pertinent.

Le coefficient alpha de Cronbach peut être calculé à l'aide de la formule suivante :

$$\alpha = \frac{N \times c}{v + (N - 1)c}$$

Où N est le nombre d'items

c est la corrélation moyenne entre les items

v est la variance moyenne

Dans le cadre du QUINOA, il faudra donc s'assurer que le coefficient alpha de Cronbach ait été calculé afin de garantir la fiabilité du test.

La fidélité test-retest doit également faire l'objet d'une préoccupation, car elle évalue la constance d'un outil de mesure dans le temps. Ainsi, entre deux passations d'un même test en un temps différent, le résultat obtenu doit être le même. La fidélité test-retest est calculée grâce à un coefficient de corrélation qui permet de mesurer l'écart des résultats d'un même test à différents instants. Il est important que ces deux passations soient réalisées par les mêmes participants, et que l'état des participants n'ait pas changé, ce qui fausserait les résultats du coefficient. Les tests doivent être administrés à un intervalle temporel assez court : DeVellis (2019) recommande d'espacer les deux passations de 2 jours à 2 semaines. Cet intervalle temporel permet que le patient ne mémorise pas les différents items et ses réponses lors de la première passation, et cela permet d'éviter un effet d'apprentissage entre les deux passations.

Vallerand, en 1989, décrit qu'un coefficient au-delà de 0,60 est satisfaisant et que plus il se rapproche de 1, plus cela signifie que le test est fiable dans le temps.

Enfin, la fidélité interjuges permet d'évaluer la constance d'un test selon la personne qui fait administrer ce test. Ainsi, dans l'idéal, il faut que plusieurs examinateurs soient présents au même moment lors d'une passation d'un même test, ou que l'examinateur s'il n'y en a qu'un filme la cession du test. A l'issue de cette passation, leurs résultats sont confrontés afin de relever les éventuelles dissonances. Cette fidélité peut être calculée en mesurant l'accord des juges, cela grâce aux coefficients de corrélation intraclasse. D'après Corbière et Larivière (2020), un coefficient intraclasse supérieur à 0,75 témoigne d'une très bonne fidélité du test. Lorsque ce résultat est inférieur à 0,40, la fidélité est considérée comme médiocre. Entre ces deux données, ils considèrent la fidélité comme acceptable.

Dans le cadre du QUINOA, le calcul de la fidélité interjuges et de la fidélité test-retest est compromis en raison des conditions de passation. Effectivement, plusieurs examinateurs ne seront pas présents aux passations, ce qui entrave la possibilité de calculer la fidélité interjuges. De plus, le test ne pourra pas être administré plusieurs fois, ce qui explique que la fidélité test-retest ne pourra pas non plus être mesurée. Cependant, nous mesurerons la cohérence interne du QUINOA.

Objectif de recherche

Les TAP sont présents chez de nombreux enfants, et aujourd’hui, il n’y a pas d’étude fiable concernant un outil de dépistage de ces troubles chez les enfants français âgés de 1 an à 6 ans. Un outil de dépistage pouvant être administré par différents professionnels de santé, permettrait d’orienter plus rapidement l’enfant et ses parents vers un orthophoniste pour une rééducation adaptée. Le QUINOA a été créé dans l’idée de devenir un outil de dépistage fiable, sur lesquels les professionnels de santé pourront s’appuyer pour repérer ces troubles le plus tôt possible.

Comme tout outil de dépistage ou de diagnostic, il faut que ce dernier ait été testé sur un échantillon représentatif de la population ciblée : c’est ce que l’on nomme l’étalonnage. L’objectif de ce travail sera de réaliser la deuxième partie de l’étalonnage du QUINOA. Pour cela, nous définirons dans un premier temps les concepts de bases, c'est-à-dire la population et les variables. Par la suite, nous recueillerons les données nécessaires à notre étude au moyen de l’administration du QUINOA puis nous les traiterons grâce à un modèle d’apprentissage : la régression logistique. Pour finir, nous calculerons, dans la mesure du possible, la fiabilité du QUINOA afin de déterminer le degré de précision de ce test.

Protocole expérimental

L'objectif de notre étude est de réaliser la deuxième partie de l'étalonnage, c'est-à-dire de recueillir les résultats des patients porteurs d'un TAP au QUINOA pensé par Barral (2018), élaboré par Grueau (2020) et pré-testé par Demonteil (2022), dans le but d'amener à la validation de cet outil de dépistage. Nous allons donc détailler le protocole expérimental mis en place pour réaliser l'étalonnage du QUINOA.

Pour ce faire, nous allons utiliser l'approche méthodologique servant aux tests diagnostiques, car bien que le QUINOA soit un test de dépistage, ces deux types de tests sont généralement construits de la même manière. Ainsi, nous allons suivre les lignes directrices de la méthodologie « Standards for Reporting of Diagnostic accuracy (STARD) » (Gedda, 2015).

1. Aspect éthique : formulaire de consentement

Un formulaire de consentement (Annexe I) a été envoyé aux professionnels de santé qui se sont portés volontaires pour administrer le QUINOA aux parents des patients inclus dans notre population d'étude. Ce formulaire devra être renseigné et signé par les titulaires de l'autorité parentale. Il est possible, en cas de séparation des parents, que les deux consentements soient recueillis sur des formulaires distincts. Ces précisions ont été indiquées aux professionnels de santé.

2. Population

Les critères d'inclusion de notre étude sont :

- Enfants âgés de 1 an à 6 ans
- Enfants porteurs d'un TAP

Comme l'a déterminé Bardousse (2019), le QUINOA est à destination des enfants de 1 an à 6 ans, et la partie de l'étalonnage sur les enfants ne présentant de TAP a été réalisée par Gendreau en 2024. Ces éléments nous ont permis de déterminer les critères d'inclusion de notre étude.

Le critère de non-inclusion de notre étude est :

- Enfants porteurs d'une pathologie à risque de causer un TAP, hormis la prématurité et le RGO qui ne sont pas exclus par ce critère.

Ce critère de non-inclusion a été mis en évidence par Barral en 2018, et nous avons choisi de le conserver dans notre étude car cette dernière vise à obtenir une distribution des scores chez des enfants tout-venant. En effet, cela risquerait de fausser les résultats de notre étude si des pathologies à risque de causer un TAP étaient prises en compte. Cependant, la prématurité et le RGO sont exclus de ce critère de non-inclusion. En France, la prématurité touche près de 60 000 nouveau-nés par an (Torchin et al. 2015), et le RGO, d'après Martigne et al (2009), aux alentours de 24% des enfants âgés de 0 à 23 mois, et autour de 7% des enfants de la tranche d'âge 2-11 ans. Un grand nombre d'enfants seraient alors exclus de l'étude si nous n'incluions pas ces critères, ainsi nous avons décidé de les conserver. D'autant

plus que 40 à 70% des enfants prématurés présenteront un jour un TAP (Le Heuzey, 2006), et que le RGO est un facteur de risque de développer un TAP.

3. Variables

La variable indépendante qui va influencer la variable dépendante, est le fait que le patient soit porteur d'un TAP ou non. Les variables dépendantes seront les réponses au questionnaire.

4. Participants

Le QUINOA a pour finalité d'être administré aux parents d'enfants porteurs d'un TAP, par le biais d'un professionnel de santé. Pour effectuer la deuxième partie de l'étalonnage du QUINOA, nous allons donc solliciter des professionnels de santé, susceptibles de prendre en soin des enfants âgés de 1 an à 6 ans porteurs d'un TAP. Pour cela, nous allons utiliser la liste des professionnels de santé ciblés par Demonteil (2022), établie à l'aide de données issues de la littérature.

Voici les professionnels sélectionnés :

- Médecins généralistes
- Pédiatres
- Médecins ou pédiatres de PMI
- Médecins scolaires
- Dentistes
- Orthodontistes
- Diététiciens
- Kinésithérapeutes
- Ergothérapeutes
- Psychomotriciens
- Orthophonistes
- Psychologues
- Psychiatres
- Pédopsychiatres
- Gastro-pédiatres
- Gastro-entérologues
- Oto-rhino-laryngologues

5. Matériel

Notre étude a pour objectif de réaliser la partie finale de l'étalonnage du QUINOA, ainsi nous allons reprendre la version consensuelle du QUINOA (Annexe II) réalisée par Demonteil (2022) après la phase de pré-test. A cette version, nous ajouterons 3 questions concernant la prise en soin de ces enfants (Annexe III), en amont de la version consensuelle du QUINOA élaborée par Demonteil (2022).

5.1. Version consensuelle du QUINOA

Le QUINOA est composé de 25 questions, dont 21 à destination des parents et 4 à renseigner par le professionnel de santé. Les informations données par le professionnel de santé concernent la date de naissance de l'enfant (ce qui nous permettra de connaître son âge), son poids et sa taille (ce qui nous permettra de calculer l'Indice de Masse Corporelle (IMC) de l'enfant), et sa tonicité oro-myo-faciale.

Les 21 questions à destination des parents sont organisées en plusieurs catégories. Les premiers aspects évalués sont les habiletés oro-motrices de l'enfant, qui font l'objet de 4 questions. Puis la sensorialité de l'enfant est estimée sur la base de 5 autres questions. Les 3 questions suivantes portent sur l'alimentation de l'enfant. Le comportement de l'enfant et de sa famille lors des repas fait l'objet d'un groupement de 7 questions. Enfin, les 2 dernières questions porteront sur l'état de santé de l'enfant.

5.2. Questions ajoutées

Aux 25 questions composant le QUINOA, nous avons choisi d'en ajouter 3, concernant une potentielle rééducation. Ces questions sont essentielles car les enfants bénéficiant d'une prise en soin ayant pour but de rééduquer leur TAP pourraient biaiser notre étude, le QUINOA ayant pour but d'être administré à des enfants tout-venant, donc sans rééducation concernant leur TAP. Ces questions porteront sur une éventuelle prise en soin du TAP, sur le type et la durée de cette prise en soin. La nouvelle version du QUINOA est disponible en annexe (Annexe IV).

Le QUINOA comporte désormais 28 questions.

5.3. Consignes de passation du QUINOA

Nous avons sensibilisé les professionnels de santé sollicités au projet au sein duquel se place notre étude : le QUINOA. Nous leur avons expliqué l'objectif final du QUINOA, ainsi que leur rôle dans ce projet, grâce à une affiche de présentation (Annexe V). Nous leur avons demandé d'administrer le questionnaire au(x) parent(s) d'un ou plusieurs enfant(s) âgé(s) de 1 an à 6 ans dont le diagnostic d'un TAP a été posé. Nous leur avons signifié que la première partie de l'étalement, auprès de patients ne présentant pas de TAP, avait déjà été effectuée. Nous leur avons également transmis le questionnaire de consentement à destination des parents, en précisant que tous les titulaires de l'autorité parentale devaient renseigner et signer ce formulaire. En cas de séparation des parents, il est possible que leurs consentements soient recueillis sur 2 formulaires distincts.

Nous avons précisé aux professionnels de santé que le questionnaire devait être renseigné par les parents en leur présence. Cette obligation fait suite aux travaux de Gendreau (2024), dans lesquels la présence du professionnel de santé lors de l'administration du questionnaire avait pour but de limiter tous les biais liés à des difficultés en lecture ou en compréhension globale du français.

Notre étude se plaçant dans la continuité des travaux de Gendreau (2024), pour que l'étalement soit le plus fiable possible lorsque le QUINOA sera publié, il faut que les conditions de passation soient les mêmes auprès des deux populations étudiées.

6. Diffusion du QUINOA

6.1. Support d'administration

Nous proposons aux professionnels de santé acceptant de participer à l'étude différents moyens d'administration du QUINOA. Nous mettons à disposition la version numérique du QUINOA, via un lien proposé dans le mail de diffusion. Ce format informatisé a été créé sur la plateforme Sphinx. Le questionnaire peut également être envoyé en version PDF pour impression, sur simple demande de leur part. Nous pouvons également faire parvenir aux professionnels une version papier du QUINOA avec une enveloppe timbrée pour le retour, s'ils le souhaitent.

6.2. Recrutement

Nous avons pour objectif de diffuser le QUINOA largement, afin d'avoir un grand nombre de participants. Pour cela, nous avons transmis le QUINOA par mail (Annexe VI.I) à un large réseau de professionnels, comprenant des médecins généralistes (8), pédiatres (4), médecins ou pédiatres de PMI (3), orthodontiste (1), diététicien (2), kinésithérapeutes (5), ergothérapeutes (4), psychomotriciens (5), orthophonistes (198), psychologues (3), pédopsychiatre (1), gastro-pédiatre (1), gastroentérologues (2), oto-rhino-laryngologue (2).

Afin d'augmenter nos chances de réponses, nous avons également diffusé le QUINOA sur certains groupes Facebook tels que :

- Ortho-néonat : un groupe privé composé d'environ 2700 personnes (orthophonistes et étudiants en orthophonie) ;
- Petites Annonces d'Orthophonie : un groupe privé constitué d'environ 19 000 personnes (orthophonistes et étudiants en orthophonie) ;
- Ortho-infos : un groupe privé réservé aux orthophonistes et étudiants en orthophonie, composé d'environ 21 500 personnes
- Orthophonie, oralité et/ou freins restrictifs : groupe de travail : un groupe privé regroupant près de 2000 orthophonistes ou étudiants en orthophonie
- Orthophonie et enfant très jeune ou sans langage : un groupe privé constitué de 3000 personnes (orthophonistes et étudiants en orthophonie)
- Troubles de l'oralité, dysoralité sensorielle, TAP : un groupe privé composé d'approximativement 3800 personnes, dont des orthophonistes et autres professionnels de santé travaillant autour des TAP.

Il était important que les groupes Facebook sur lesquels a été diffusé le questionnaire soient privés et consacrés aux professionnels de santé, afin de garantir une fiabilité des réponses.

Nous avons renvoyé le QUINOA aux professionnels ne nous ayant pas répondu : une première relance a été adressée au mois de décembre 2024 (Annexe VI.II) et une seconde au mois de janvier 2025 (Annexe VI.III).

7. Méthode d'analyse des résultats

Notre objectif de recherche est de déterminer des seuils d'alerte aux réponses du QUINOA, en fonction des résultats des sujets porteurs d'un TAP et des sujets ne présentant pas de TAP. Pour cela, nous allons adopter une approche comparative, en confrontant les résultats recueillis par Gendreau (2024) et ceux que nous allons collecter.

Gendreau (2024) avait déterminé, grâce à une étude sur la population non-porteuse d'un TAP, que le seuil d'alerte du QUINOA se situait entre 22,11 et 27,51 points avec un intervalle de confiance à 95% grâce au calcul du z-score. Le mode de calcul des quantiles à 95% donnait quant à lui un seuil d'alerte à 22 points. Pour ce faire, elle avait transformé les résultats qualitatifs en résultats quantitatifs (Annexe VII). Notre étude se situant dans la continuité de la sienne, nous allons reprendre son score ainsi que l'échelle utilisée lors de son travail.

Nous allons comparer le score déterminé par Gendreau (2024) avec les résultats obtenus lors de cette phase d'étalonnage, mais nous souhaitons également traiter les données grâce au modèle de régression logistique, qui permet de minimiser le pourcentage d'erreur. Ainsi, nous comparerons nos résultats avec l'approche de Gendreau (2024) qui est plus classique.

La date de clôture de la diffusion du QUINOA a été fixée au 31 janvier 2025. Une fois que toutes les données auront été récoltées, nous allons les prétraiter grâce au logiciel Python. Cette étape de prétraitement aura pour objectif de transformer les données brutes en données chiffrées. Nous avons rédigé des lignes de codes (Annexe VIII) sur Python, permettant d'attribuer les scores définis par Gendreau (2024). Cela conduira à transformer le résultat qualitatif en résultat quantitatif d'après l'échelle de Gendreau (2024) en minimisant le risque d'erreur. Ensuite, l'objectif sera de traiter les données grâce à la régression logistique, pour parvenir à notre objectif de recherche, qui est de déterminer si un enfant tout-venant est potentiellement porteur d'un TAP ou non. Pour cela, nous avons rédigé un script ayant pour objectif d'entraîner les modèles de régression logistique (Annexe IX) que nous programmerons sur Python. Ce script nous permettra également de calculer la sensibilité et la spécificité des outils entraînés, ainsi que de réaliser la régularisation L1, grâce à laquelle nous pourrons supprimer les questions qui ont les plus faibles poids dans la décision d'un TAP, et donc de réduire le temps de passation du QUINOA.

En parallèle, nous allons faire une ACP, ayant pour but de synthétiser les données et d'avoir une image globale des données issues du questionnaire. Le script pour ce faire est détaillé dans les annexes (Annexe X).

Par la suite, nous comparerons graphiquement les deux méthodes utilisées : celle de Gendreau (2024) et celle par régression linéaire. Nous calculerons également les métriques de ces outils de décision. Pour cela, nous allons générer des diagrammes de dispersion : un premier ayant pour composante principale la régression logistique, et un second ayant pour composante principale le score de Gendreau (2024). Pour comparer les deux outils, nous avons également tracer les deux courbes ROC grâce à un autre script à introduire après le script de pré-traitement (Annexe XI.)

L'objectif final du QUINOA est de déterminer rapidement si un enfant tout-venant est potentiellement porteur d'un TAP. Ainsi, nous devrons créer une page internet ou un logiciel, sur lesquels les médecins généralistes faisant passer le QUINOA rentreront les résultats. A l'issue de la passation, le site internet ou le logiciel donnera la décision finale sur la possibilité que l'enfant soit porteur d'un TAP ou non. Cette étape ne sera pas réalisée dans le cadre de notre travail et devra faire l'objet d'une prochaine étude.

Présentation des résultats

L'objectif de notre étude était donc de recueillir les réponses du QUINOA concernant des enfants âgés de 1 an à 6 ans, porteurs d'un TAP. Nous allons donc présenter les résultats obtenus.

1. Questionnaires inclus dans l'étude

Au total, nous avons recueilli 22 questionnaires : 19 sur le logiciel Sphinx, 2 ont été envoyés par voie postale et 1 nous a été transmis par mail. Parmi ceux-là, nous avons exclu un enfant qui n'avait pas l'âge requis (9 ans) et un enfant présentant un trouble du spectre autistique alors qu'il s'agissait d'un critère d'exclusion. Au total, nous avons donc 20 questionnaires qui pourront être analysés.

2. Participants

Les 20 participants dont nous allons traiter les données sont âgés de 1 an 5 mois à 5 ans 9 mois.

Nous allons désormais répertorier ces patients selon les tranches d'âges déterminées : dans la première tranche d'âge incluant les enfants de 1 an à 2 ans 11 mois, nous comptons 9 patients. Dans la deuxième tranche d'âge englobant les enfants de 3 ans à 4 ans 11 mois, nous avons recensé 8 patients. Enfin, dans la dernière tranche d'âge qui s'étend de 5 ans à 6 ans, les patients sont au nombre de 3.

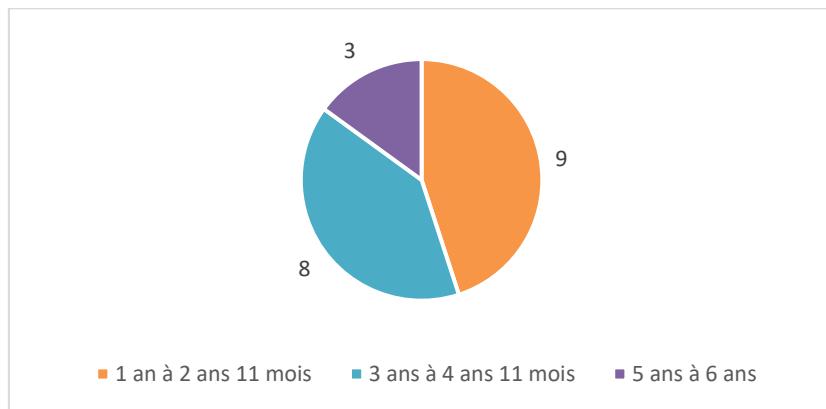


Figure 2 : Répartition des participants en fonction de l'âge

3. Temps de passation

Concernant le temps de passation, nous avons seulement les données des questionnaires remplis sur Sphinx.

Une donnée est largement supérieure aux autres (2487 secondes), il s'agit donc d'une valeur extrême. Pour le traitement des données, nous avons choisi d'exclure cette valeur car elle

n'est pas représentative de la population. Le temps moyen de passation est 483 secondes, ce qui correspond à 8 minutes et 3 secondes. Le temps maximal de passation est de 950 secondes, autrement dit environ 16 minutes. Le temps minimal de passation est de 128 secondes, ce qui correspond à peu près à 2 minutes.

Nous avons établi un graphique récapitulant les données recueillies concernant le temps de passation :

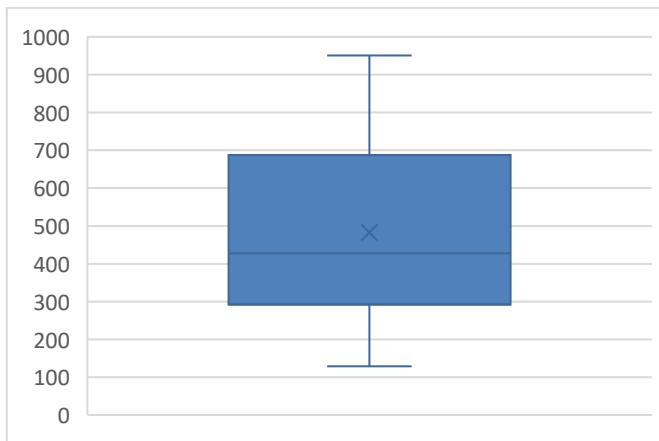


Figure 3 : Temps de passation du questionnaire, en secondes

4. Résultats obtenus

Les résultats obtenus à chaque question sont détaillés en annexe (Annexe XIV).

5. Calcul du score

Pour calculer le score total de chaque questionnaire, nous allons reprendre l'échelle utilisée par Gendreau (2024), ce qui nous permettra de conserver une cohérence entre la première partie de l'étalonnage et la seconde. Comme elle l'avait déterminé, certaines questions ont une répartition des points sur un barème de 0 à 1 (question 5), ou de 0 à 2 (question 6), ou de 0 à 3 (questions 7, 19, 12, 18 et 22) ou encore de 0 à 4 (questions 8, 9, 11, 13, 14, 15, 16, 17, 19, 20, 21, 23, 24, 25, 26). Dans son échelle, Gendreau (2024) note toutes les questions sur 4 points au total. Ainsi, nous avons fait la conversion pour chaque réponse. Tous ces scores sont détaillés dans les annexes (Annexe VII).

6. Distribution des scores obtenus

6.1. Statistiques globales

Parmi les 20 questionnaires recueillis, les scores totaux sont compris entre 21,66 et 49,66. Ils peuvent être récapitulés dans ce graphique :

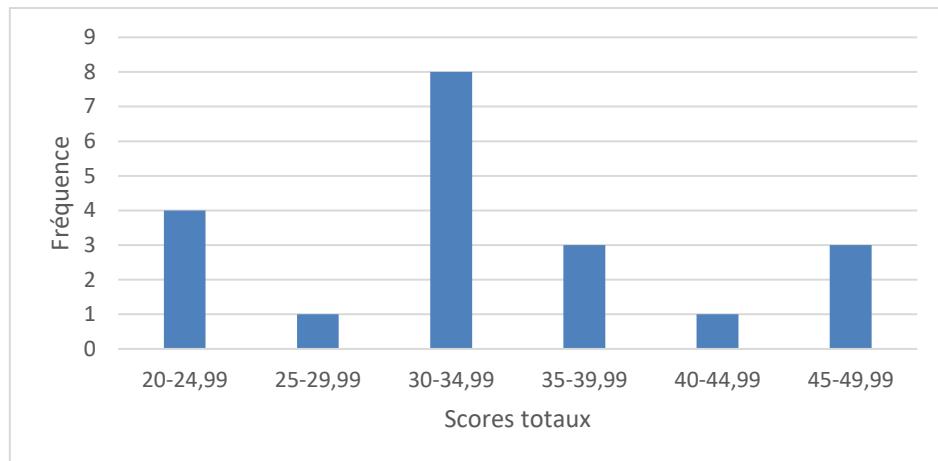


Figure 4 : Distribution des scores obtenus

Grâce au logiciel R, nous avons calculé la moyenne des scores totaux obtenus, qui est équivalente à 34,21 points.

6.2. ACP

Pour rappel, l'ACP permet de visualiser l'ensemble des données en réduisant les dimensions. L'abscisse et l'ordonnée correspondent aux deux composantes principales.

Nous avons généré l'ACP avec les résultats que nous avons collectés et ceux obtenus par Gendreau en 2024 :

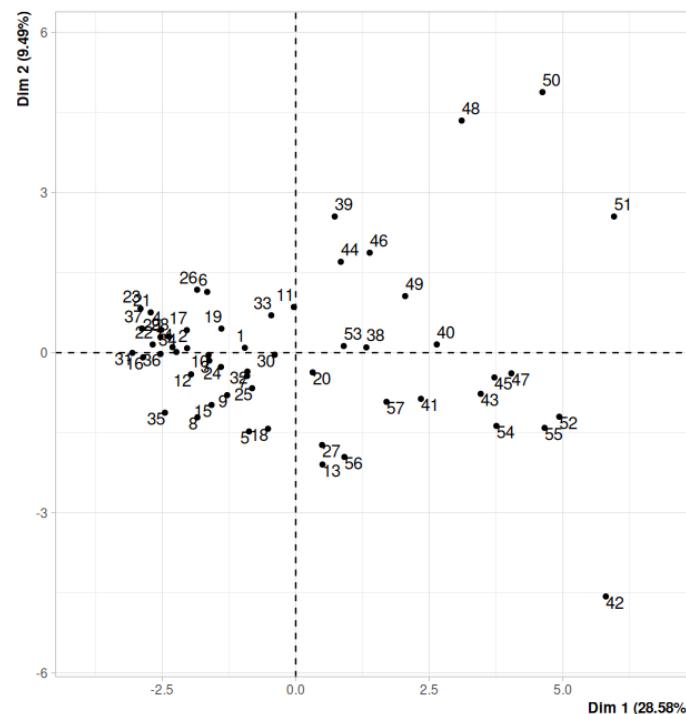


Figure 5 : Premier graphique de l'ACP des scores obtenus aux deux phases d'étalonnage du QUINOA

Les points numérotés 0 à 36 correspondent aux résultats obtenus par les enfants non-porteurs d'un TAP et ceux numérotés de 37 à 56 correspondent aux résultats recueillis auprès des

enfants porteurs d'un TAP. Les pourcentages en abscisse et en ordonnée mettent en évidence le fait que 38,07% de l'information sous-tendue par ce graphique a été conservée dans l'ACP.

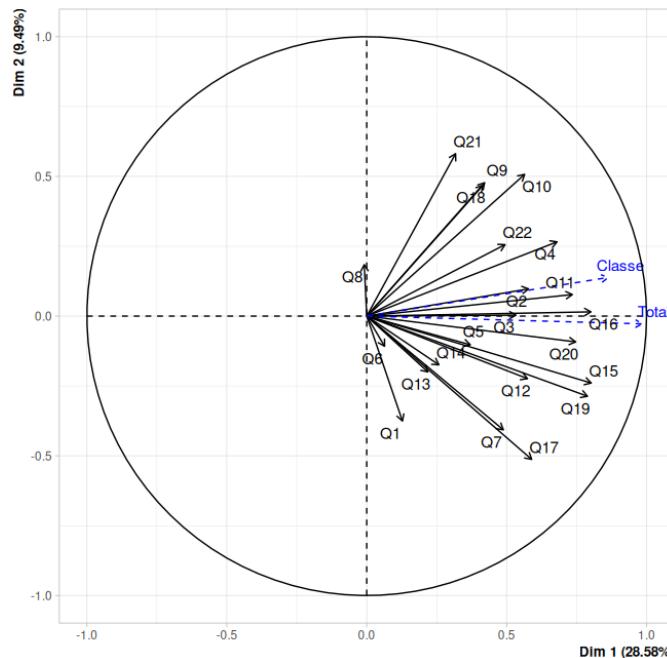


Figure 6 : Deuxième graphique de l'ACP des scores obtenus aux deux phases d'étalonnage du QUINOA

La classe correspond au fait que l'enfant soit porteur ou non d'un TAP. Le total représente le score obtenu au QUINOA. Les différents vecteurs correspondent aux questions du QUINOA, les variables. On voit que la classe et le total tendent dans la même direction, ce qui signifie que ces deux éléments sont possiblement corrélés.

7. Détection d'un patient susceptible de présenter un TAP

L'intérêt de notre étude est de trouver un moyen de déterminer si un enfant, après que ses parents ont renseigné le QUINOA, est susceptible ou non de présenter un TAP. A présent, nous allons envisager les différentes manières de procéder.

7.1. Les méthodes de Gendreau (2024) : le seuil d'alerte

Gendreau (2024) a déterminé des seuils d'alerte qui correspondent aux scores à partir desquels un enfant est potentiellement porteur d'un TAP. Ces seuils d'alerte ont été déterminés selon deux méthodes différentes : le calcul du z-score et les quantiles. Nous allons détailler ces méthodes et les mettre en lien avec les scores obtenus à notre étalonnage.

7.1.1. Le seuil d'alerte déterminé à partir du z-score

Gendreau a déterminé un seuil d'alerte à partir du calcul z-score :

Moyenne + 2 x écart-type

Avec un intervalle de confiance à 95%, elle a déterminé que le score seuil se situe entre 22,11 et 27,51. Nous souhaitons que le QUINOA soit le plus sensible possible, c'est-à-dire qu'il dépiste tous les enfants porteurs d'un TAP sans exception, c'est pour cela que nous conservons pour notre étude le score le plus bas, soit 22,11.

7.1.2. Le seuil d'alerte déterminé à partir des quantiles

Gendreau (2024) a déterminé un score seuil à partir des quantiles, car la distribution des scores du QUINOA ne suit pas une loi normale. D'après cette méthode, le score seuil s'élève à 22 avec un intervalle de confiance à 95%.

7.2. Une autre méthode : la régression logistique

Comme nous l'avons vu précédemment, le modèle de régression logistique est un outil de décision qui permet de déterminer, selon les résultats obtenus au questionnaire, l'appartenance à une classe (Hastie et al., 2009). Dans notre cas, les deux groupes possibles sont : « l'enfant est potentiellement porteur d'un TAP » et « l'enfant n'est pas porteur d'un TAP ».

Nous allons utiliser le modèle de régression logistique qui permet de prédire une variable dépendante à partir d'une ou plusieurs variables indépendantes. Nous rappelons que dans notre étude, la variable indépendante est le fait que le patient soit porteur ou non d'un TAP et les variables dépendantes sont les réponses au questionnaire.

La régression logistique met en évidence un coefficient ou poids pour chaque question, visible dans le calcul du score présenté précédemment, que nous rappelons :

$$S = \langle W, X \rangle + b$$

Où $\langle \cdot, \cdot \rangle$ désignent le produit scalaire usuel

W désigne le vecteur des poids associés à chaque question

X est le vecteur contenant les réponses à chacune des questions

b est un nombre appelé le biais.

Ce score est déterminé grâce aux réponses obtenues dans notre phase d'étalonnage. Le poids est déterminé quant à lui lors du processus d'apprentissage sur le jeu de données d'entraînement, détaillé en annexe (Annexe XV). La régression logistique fait apparaître différents poids : des positifs, des négatifs, des plus petits et des plus grands. Un poids qui a un score nul correspond à une question qui n'intervient pas dans la décision finale, alors qu'un poids très positif indique une question très influente.

Contrairement à la régression logistique, Gendreau (2024) en établissant son échelle de conversion de résultats qualitatifs en résultats quantitatifs, avait fait le choix de donner la même importance à chaque question. Ainsi, toutes les questions avaient le même poids dans son étude.

Après avoir saisi les scripts dans Python comme le prévoyait le protocole expérimental, nous obtenons un score de 100% en sensibilité, spécificité et précision sur le jeu de données de test, ce qui signifie que l'apprentissage supervisé fonctionne et ne donne jamais de réponse erronées sur le jeu test.

Ainsi, cette méthode présente un seuil d'alerte formulé par le calcul du score et permet également d'évaluer la probabilité d'être porteur d'un TAP en fonction des réponses aux questions, grâce à la formule ci-dessous :

$$P(TAP) = \frac{\exp^s}{1 + \exp^s}$$

Où P(TAP) est la probabilité que l'individu soit classé comme TAP pour le modèle de régression logistique

\exp est la fonction exponentielle

s est le score obtenu

8. Les métriques des outils de décision

Pour calculer les caractéristiques des deux outils de décision, nous avons extrait 9 cas de la population totale (comprenant les deux phases d'étalonnage) : il s'agit du jeu de données de test. Ces valeurs ont été sélectionnées aléatoirement afin de mesurer les performances de la régression logistique sur des nouvelles données, suite au traitement du jeu de données d'entraînement. Pour pouvoir comparer la méthode de la régression logistique et celle du score-seuil, nous avons mesuré leurs performances.

Les valeurs relevées sont les suivantes :

k	0	1	2	3	4	5	6	7	8
Probabilités issues du modèle de régression logistique que le patient soit porteur d'un TAP	0	0,15	0,98	0	0,99	0,06	0,15	0	0
Scores de Gendreau (2024)	7,66	21	33,99	6,99	39,98	21,32	20,66	9,99	9,32
Vrai y	0	0	1	0	1	0	0	0	0

Tableau 2 : Comparaison des résultats entre la régression logistique et la méthode de Gendreau (2024) sur le jeu test

Où k représente le numéro du cas étudié, attribué au hasard

Vrai y correspond au fait que l'enfant ait un diagnostic de TAP ou pas :

- $y = 0$: l'enfant n'est pas porteur d'un TAP
- $y = 1$: l'enfant est porteur d'un TAP.

Les probabilités mises en évidence par le modèle de régression logistique montrent que les participants se situant très proche de 1 ($k = 2$ et $k = 4$), sont réellement porteurs d'un TAP ($y = 1$). Les participants ayant obtenu une probabilité très faible, ne sont effectivement pas porteurs d'un TAP ($y = 0$).

Nous pouvons donc établir le tableau suivant à partir de l'une ou de l'autre des méthodes :

	Enfant non-porteur d'un TAP	Enfant porteur d'un TAP	Total
$y = 0$	7 = vrai-négatifs	0 = faux-négatif	7 = total « $y = 0$ »
$y = 1$	0 = faux-positif	2 = vrai-positifs	2 = total « $y = 1$ »
Total	7 = total « enfants non-porteurs d'un TAP »	2 = total « enfants porteurs d'un TAP »	9 = total

Tableau 3 : Résumé des résultats d'après la méthode de Gendreau (2024) et d'après le modèle de la régression logistique sur le jeu test

8.1. Le seuil d'alerte selon Gendreau (2024)

Pour rappel, Gendreau (2024) avait fixé le seuil d'alerte à 22 points.

Nous avons calculé la sensibilité, la spécificité, la précision et les VPP et VPN sur le jeu de données test. Ces 4 valeurs sont égales à 1.

8.2. La régression logistique

Les métriques permettant de détailler la précision d'un outil de décision ont été calculées par la régression logistique (Annexe IX). La sensibilité et la spécificité ont un taux de 100%, tout comme la précision, la VPP et la VPN, sur le jeu test ainsi que sur le jeu d'entraînement.

8.3. Visualisation des deux détecteurs

Nous avons réalisé des graphiques permettant de montrer les différences entre les deux outils de détection. Sur ces graphiques, les réponses obtenues lors de la première phase d'étalonnage par Gendreau (2024) sont schématisées en bleu, et celles recueillies lors de la deuxième phase d'étalonnage sur une population porteuse d'un TAP, en rouge. Sur l'ACP il est possible de visualiser le score de Gendreau (2024) et le score de la régression logistique mais cela n'est pas pertinent car ces directions, bien qu'étant visuellement très corrélées avec l'axe des abscisses (composante principale 1), ne permettent pas de visualiser de différence claire entre ces deux directions sur l'ACP.

Nous avons donc décidé, afin de produire un graphique plus lisible, de remplacer la composante principale 1 par l'un des scores et d'utiliser comme axe des ordonnées une

« composante principale 2 » obtenue en maximisant la variance du nuage de points étant donné la direction fixée par le score en abscisse.

Le premier graphique a été conçu du point de vue du score généré par Gendreau, en 2024 :

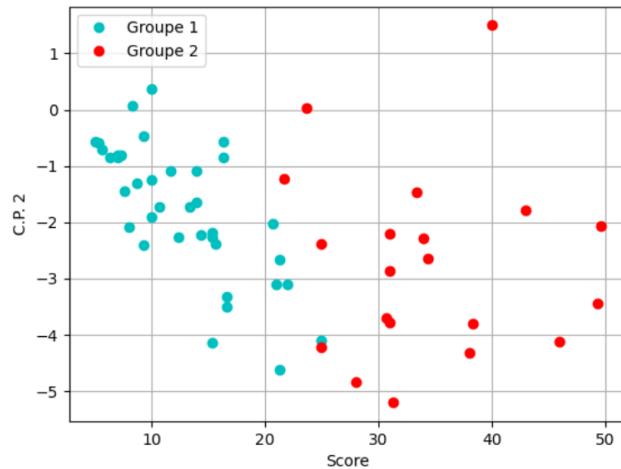


Figure 7 : Graphique du point de vue du score de Gendreau (2024)

Ce graphique met en évidence que les enfants porteurs d'un TAP ont un score plus élevé, en moyenne, que les enfants ne présentant pas ce trouble. Il y a visuellement une distinction claire entre les deux groupes : les enfants porteurs d'un TAP sont plutôt placés à droite du graphique tandis que les enfants non-porteurs d'un TAP sont sur la gauche du graphique. Cependant, on note que certains scores des deux groupes sont proches, voir égaux et qu'il n'existe pas de séparation nette entre les deux groupes.

Le deuxième graphique que nous avons généré se place du point de vue de la régression logistique que nous avons créé :

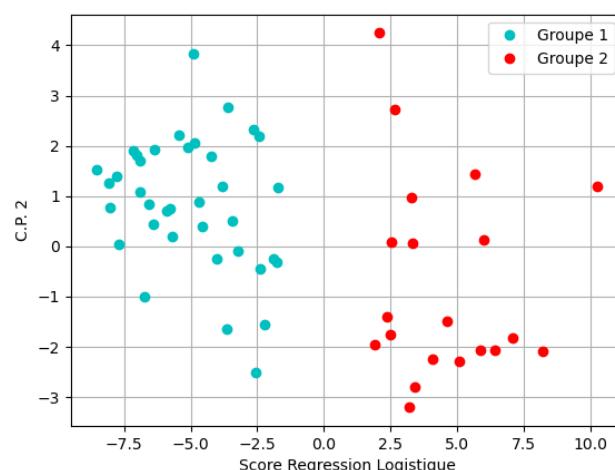


Figure 8 : Graphique du point de vue de la régression logistique

Sur ce graphique, nous remarquons à nouveau que les enfants porteurs d'un TAP sont placés sur la droite du graphique, ils ont donc un score plus élevé que les enfants ne présentant pas de TAP qui se situent à gauche du graphique. Cependant, il apparaît une nette séparation entre les deux groupes contrairement au premier graphique.

8.4. Courbes ROC

Nous avons également tracé les courbes ROC des deux outils de décision. Elles sont représentées sur le même graphique : la courbe ROC correspondant au seuil d'alerte déterminé par Gendreau (2024) est schématisée en orange et la courbe ROC du modèle de régression logistique du QUINOA en bleu :

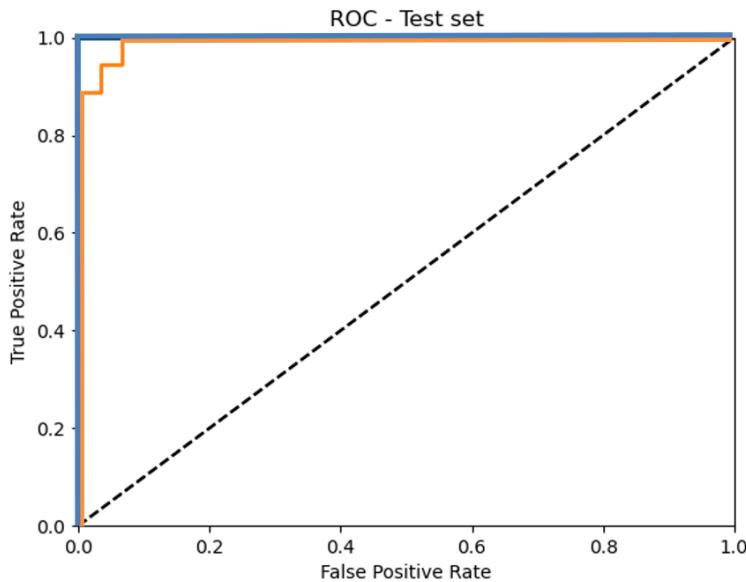


Figure 9 : Courbes ROC du modèle de Gendreau (2024) et du modèle de régression logistique

Ce graphique montre que la courbe ROC du modèle de régression logistique est supérieure à celle du modèle de Gendreau (2024), et donc que la sensibilité et l'antispécificité sont meilleures sur le modèle de régression logistique. La courbe ROC du modèle de Gendreau (2024), bien qu'en dessous de celle du modèle de régression logistique, est cependant très haute, ce qui montre que cette méthode est aussi très fiable.

9. Régularisation l1

Afin de déterminer les questions les moins pertinentes dans le processus de décision, nous avons utilisé une régularisation l1. Cette régularisation va nous permettre d'annuler le poids de ces questions. Ce tri s'effectue comme dit précédemment, par la sélection d'un coefficient de régularisation afin d'obtenir un modèle possédant un biais et une variance corrects. Nous avons donc décidé de sélectionner une valeur de ce coefficient de manière à ce que les performances de l'outil ne soient pas dégradées par cette régularisation (Annexe IX).

Ici, nous avons choisi 0,7 comme coefficient de régularisation. Il s'agit du point où le biais et la variance sont les plus faibles. La régularisation l1 nous a permis de supprimer 13 questions :

- Question 5 : « Lèvres/langue/joues/mâchoire de l'enfant sont-elles toniques ? »
- Question 8 : « L'enfant accumule-t-il de la nourriture dans ses joues et/ou la recrache après l'avoir stockée ? »
- Question 9 : « En général, le repas de l'enfant dure en moyenne : »
- Question 11 : « A quelle fréquence l'enfant : - refuse qu'on lui touche le visage ou le corps (caresses, bain, câlins) ou - refuse le contact des mains, des pieds ou du corps avec certaines matières (aliments, textiles, herbe, sable, pâte à modeler...) ? »
- Question 12 : « L'enfant porte-t-il (ou a-t-il porté) les objets ou ses doigts à la bouche en dehors des périodes de poussées dentaires ? »
- Question 14 : « A quelle fréquence l'enfant a-t-il des haut-le-cœur lors de l'introduction d'un aliment (en dehors des aliments favoris ou acceptés depuis toujours) ou d'un objet en particulier dans sa bouche ? »
- Question 16 : « L'enfant refuse-t-il de goûter des aliments nouveaux, choisit les aliments qu'il peut manger et trie son assiette ? »
- Question 17 : « L'enfant a-t-il une préférence pour les aliments au goût prononcé ? (Par exemple : cornichons, moutarde ...) ou L'enfant prend-il une grande quantité d'aliments en bouche ? »
- Question 18 : « L'enfant manifeste-t-il sa faim ? »
- Question 19 : « L'enfant redoute-t-il les moments de repas, le passage à table se fait-il difficilement ? »
- Question 21 : « L'enfant réagit-il différemment (refus de manger...) selon la personne avec qui il mange ? »
- Question 22 : « L'enfant a-t-il peur d'avaler et/ou de s'étouffer ? »
- Question 25 : « L'enfant fait-il des fausses routes (la nourriture passe de travers) ? »

Le coefficient de régularisation choisi élimine plus de la moitié des questions.

Nous avons calculé les métriques de performance du modèle de régression logistique avec la régularisation l1, ce qui met en évidence une spécificité et une sensibilité sur le jeu d'entraînement et le jeu test égales à 1.

10. L'équilibre biais/variance du modèle de régression logistique

Dans notre expérience, il n'y a pas de sur-apprentissage du modèle, car les mesures de performance sur le jeu de données test sont égales aux mesures de performance sur le jeu de données d'entraînement. Ainsi, l'équilibre biais/variance est bon, ce qui signifie que le biais et la variance sont relativement bas.

11. Les aspects psychométriques du QUINOA

11.1. La cohérence interne

Comme nous l'avons présenté précédemment, l'alpha de Cronbach permet de mesurer la corrélation entre les questions d'un test, dans l'objectif de révéler la fiabilité du test. Il se calcule grâce à la formule suivante :

$$\alpha = \frac{N \times c}{v + (N - 1)c}$$

Où N est le nombre d'items

c est la corrélation moyenne entre les items

v est la variance moyenne

L'alpha de Cronbach d'après le modèle de régression logistique utilisé, est égal à :

$$\alpha = \frac{22}{21} \frac{147^2 - 29^2}{147^2}$$

$$\alpha = 0,96$$

L'alpha de Cronbach est donc égal à 0,96.

11.2. La fiabilité test-retest

La fiabilité test-retest n'a pas pu être calculée, car le QUINOA n'a pas été proposé plusieurs fois au même enfant.

11.3. La fidélité interjuges

La fidélité interjuges n'a pas pu être calculée, car lors de chaque administration du QUINOA, un seul examinateur était présent.

Discussion

L'objectif de notre recherche était de réaliser l'étalonnage du QUINOA sur des patients porteurs d'un TAP, afin de déterminer une norme, pour laquelle un enfant tout-venant est susceptible d'être porteur d'un TAP ou pas.

1. Questions relatives à une prise en soin

Nous avions ajouté des questions concernant la prise en soin des enfants porteurs d'un TAP. Nous pensions qu'elles allaient avoir un impact sur les réponses au questionnaire. Cependant, nous avons constaté que les réponses étaient variées et que le critère de la prise en soin n'influait pas sur ces réponses. Ainsi, nous avons décidé de ne pas tenir compte de ces questions dans notre analyse.

2. Comparaison des méthodes

2.1. Synthèse des données par l'ACP

Concernant l'ACP générée sur la totalité des individus inclus dans l'étalonnage, soit les figures 5 et 6, nous avions remarqué sur les deux axes principaux (l'abscisse et l'ordonnée) que 38% de l'information avait été conservée. On peut donc penser que l'ACP engendre une perte d'information conséquente cependant il faut tenir compte de l'impossibilité de réaliser un graphique avec toutes les composantes de cette situation. Sur le graphique des variables de l'ACP, soit la figure 6, nous pouvons observer que tous les vecteurs sont orientés plus ou moins vers la droite du graphique, dans la même direction que la classe et le score. Cela montre que les questions tendent toutes vers le même but : déterminer si un enfant est potentiellement porteur d'un TAP.

Sur la première ACP (figure 5), on constate que tous les enfants porteurs de TAP sont à droite alors que tous les autres sont à gauche. On a donc, sans avoir utiliser un outil de détection, une séparation visuelle entre les deux groupes, ce qui indique que le recrutement a été opérant.

2.2. Discussion des méthodes de Gendreau (2024)

En regard des résultats obtenus à la deuxième partie de l'étalonnage, soit sur la population porteuse d'un TAP, nous faisons l'hypothèse que tous les scores obtenus dans notre étude sont supérieurs aux seuils d'alerte déterminés par Gendreau (2024).

Les calculs des caractéristiques de l'échelle de Gendreau (2024) montrent que celle-ci est très sensible, très spécifique et que les VPP et VPN sont très hautes, sur le jeu de données test. Cela signifie que le test est très précis pour décider de la présence d'un TAP. Cependant, ces résultats ne sont pas parfaits, car nous avons pu remarquer que sur la courbe ROC, les métriques ne sont pas au maximum.

De plus, nous avons pu voir dans la première ACP que certains scores d'enfants non-porteurs d'un TAP et d'enfants porteurs d'un TAP se chevauchaient. Si nous étudions toutes les données, nous voyons que certains enfants porteurs d'un TAP ont un score en dessous de 22, et certains enfants non-porteurs d'un TAP ont un score au-dessus de 22. La population

étudiée est donc soit trop petite soit pas assez générale pour pouvoir conclure. Un échantillon plus important permettrait de pouvoir faire davantage de généralités.

De plus, Gendreau, en 2024, avait déjà discuté sa méthode, à propos du fait que certaines questions avaient reçu des réponses qui se distribuaient de façon proportionnelle entre les quatre options proposées. Elle avait déjà questionné le fait de considérer comme un signe d'alerte, un comportement qui est fréquent au sein d'une population saine. Cela nous amène à questionner l'utilité de l'échelle de Gendreau (2024), et à envisager l'abandon de cette échelle au profit d'autres modèles d'apprentissage qui décideraient de l'échelle à utiliser pour chaque question à partir des données brutes. Nous ne traiterons pas cette question dans notre étude.

2.3. Discussion du modèle de régression logistique

Comme nous l'avons vu précédemment, le fait d'utiliser un apprentissage supervisé permet de préciser la formule du score, en utilisant des poids déterminés par l'apprentissage. Cette formule plus générale du score permet de traiter les données avec un risque moindre d'erreurs dans la classification.

Le modèle de régression logistique, contrairement à la méthode des seuils utilisée par Gendreau (2024), met en évidence des coefficients différents pour chaque question. Cela conduit à déterminer les questions les plus sensibles pour dépister un TAP. Ainsi, certaines questions sont apparues plus importantes que d'autres dans la détection d'un TAP, comme les questions 6 et 15 qui ont un plus gros coefficient (respectivement 0,902 et 0,857) que les questions 5 et 19 aux coefficients plus faibles (respectivement $3,22 \times 10^{-4}$ et 0,114). En effet, les questions « Lèvres/langue/joues/mâchoire de l'enfant sont-elles toniques ? » et « L'enfant redoute-t-il les moments de repas, le passage à table se fait-il difficilement ? » ont donc peu d'importance dans la décision du TAP selon notre modèle. Par ailleurs, certaines questions se révèlent avoir un poids négatif, telles que les questions 10 et 18 (respectivement -0,110 et $-4,26 \times 10^{-2}$). Cela montre que, plus le score obtenu à ces questions est élevé, moins la personne a de probabilité d'être porteuse d'un TAP. La question 18, « L'enfant manifeste-t-il sa faim ? » a un poids très proche de zéro, ce qui signifie qu'elle entre peu en compte dans la décision d'un TAP. Cependant, la question 10, « Estimez-vous que votre enfant est plus "casse-cou" que les autres enfants de son âge ? » a un poids négatif important, ce qui nous interroge le lien entre le fait qu'un enfant soit « casse-cou » et la présence d'un TAP. D'après notre modèle, le fait qu'un enfant soit considéré comme « beaucoup plus "casse-cou" » ou « plus prudent » (qui sont les réponses présentant le plus de points à cette question) n'évoque pas un TAP. Cela nous conduit à remettre en question le choix des items effectué par Grueau en 2020. L'échantillonnage restreint utilisé pour effectuer l'étalonnage pourrait aussi expliquer le poids négatif de certaines questions.

Ces poids ont été soumis à la régularisation l1. Ainsi, les questions 6 et 15 qui possédaient un poids important ont été conservées, tandis que les questions 5, 18 et 19 qui avaient un poids proche de zéro ont été supprimées. Quant à la question 10, elle a été conservée bien que d'après notre modèle, un score élevé à cette question oriente vers l'absence d'un TAP.

Nous avons calculé les métriques de performance pour notre modèle et nous avons pu voir que ces métriques avaient la plus grande valeur possible, ce qui témoigne d'une précision infaillible. Ces métriques ont été calculées sur le jeu de données de test mettant en évidence que le modèle n'était pas en sur-apprentissage. Cependant, ces scores sont à nuancer avec le fait que l'étalonnage ait été effectué sur une population réduite et trop spécifique.

2.4. Comparaison des deux méthodes

Pour comparer les deux méthodes, nous allons décrypter les graphiques générés. Le premier graphique du point de vue du score élaboré par Gendreau (2024) (figure 7), montrait qu'il n'y avait pas de séparation nette entre les enfants issus de la première phase de l'étalonnage et ceux issus de la deuxième. Gendreau, en 2024, avait précisé lors de la diffusion du QUINOA pour la première phase de l'étalonnage, que les enfants prenant part à l'étude « ne devaient pas présenter de TAP, diagnostiqué ou suspecté, ni de pathologie à risque de créer un TAP ». Ainsi, le score seuil de Gendreau (2024) s'est révélé peu fiable pour distinguer les deux classes, puisque les scores des enfants porteurs d'un TAP et de ceux ne présentant pas ce trouble se chevauchent.

Le graphique qui se plaçait sous le regard de la régression logistique (figure 8), montre quant à lui une nette séparation entre les enfants porteurs du trouble et ceux qui ne le sont pas. Ainsi la régression logistique apparaît plus fiable que le score seuil élaboré par Gendreau (2024) pour déterminer si un patient tout-venant est porteur d'un TAP ou non. Effectivement, ce graphique n'émet aucun doute sur la classe dans laquelle tous les participants se placent. Ces résultats sont à nuancer avec la probabilité que l'échantillon manque sans doute de variabilité. Effectivement, l'échantillon réduit n'est pas forcément représentatif de la population ciblée, car il existe une plus grande variabilité dans la population globale que dans l'échantillon.

Nous pouvons donc conclure que la régression logistique permet une meilleure séparation des catégories d'enfants porteurs d'un TAP et d'enfants non-porteurs d'un TAP que le score généré par Gendreau (2024).

2.5. Les aspects psychométriques du QUINOA

Comme nous l'avons vu précédemment, l'alpha de Cronbach est égal à 0,96 dans le cadre de la régression logistique. Ce résultat montre que certaines questions du QUINOA sont considérées comme redondantes (Streiner et Norman, 1995). Cette redondance est signe que les questions sont pertinentes et orientent dans la même direction : celle d'être porteur d'un TAP. Ainsi, cela n'entache en rien la fiabilité du test.

3. Biais méthodologiques et limites de l'étude

Nous avions pour objectif de réaliser la deuxième partie de l'étalonnage du QUINOA auprès d'enfants porteurs d'un TAP. Pour cela, nous avons choisi d'utiliser un modèle de régression logistique, afin de minimiser le risque d'erreur. Nous avons malgré cela rencontré des biais méthodologiques et des limites à cette étude.

Le premier biais méthodologique concerne le diagnostic de TAP qui devait être posé en amont, pour pouvoir inclure l'enfant à notre étude. Le diagnostic de ce trouble étant complexe à poser, car il découle d'observations cliniques, nous pouvons supposer que le diagnostic reste dépendant du professionnel de santé qui le pose. Il s'en suit qu'il est possible que tous les participants de la deuxième phase de l'étalonnage n'aient pas été porteurs d'un TAP. Il aurait fallu vérifier tous les critères de diagnostic du TAP, par nous-même, auprès de chaque participant. Cependant, cette démarche supplémentaire aurait probablement découragé des parents de participer à l'étude. De plus, cela pouvant remettre en question la compétence du professionnel de santé ayant posé le diagnostic, il aurait été délicat d'effectuer cette étape.

Il est apparu un autre biais méthodologique relatif aux conditions de passation. En effet, nous n'avions pas précisé que les deux parents devaient être présents lors de la passation du QUINOA. Sachant que les deux parents n'ont certainement pas les mêmes ressentis ni les mêmes observations à l'égard de leur enfant, il y aurait certainement eu de légères différences avec les résultats obtenus si les deux parents avaient été présents à chaque passation. Nous n'avions pas souhaité inclure cette consigne de passation du questionnaire, redoutant de restreindre encore le nombre de participations.

Enfin, le dernier biais méthodologique, toujours relatif aux conditions de passation, concerne la fidélité inter-juges : les professionnels de santé ont réalisé la passations librement, sans contrôle, s'affranchissant possiblement des consignes données.

Notre étude a rencontré une limite du fait de l'échantillon restreint recueilli (20 patients) et sans doute peu représentatif de la population globale des enfants âgés de 1 an à 6 ans porteurs d'un TAP. Il est donc primordial de nuancer les résultats obtenus, notamment à propos des poids générés par la régression logistique, mais aussi de la sensibilité, la spécificité, la précision, les valeurs prédictives, qui s'ont amenés à évoluer lors de la poursuite l'étalonnage. Avec un échantillon plus important, les métriques du QUINOA seraient sans doute moins optimales, mais cela permettrait que l'échantillon étudie des profils plus variés, et augmenterait ainsi la fiabilité du QUINOA. Pour conclure, l'outil de décision pourrait être encore amélioré en poursuivant l'étalonnage. Concernant cette limite, nous rappelons que le QUINOA était initialement réparti en 3 tranches d'âge :

- 1 an à 2 ans 11 mois
- 3 ans à 4 ans 11 mois
- 5 ans à 6 ans

Nous avons choisi de ne pas traiter les données selon ces tranches d'âges, au vu du nombre de participations trop réduit dans chaque classe d'âge : 9 enfants étaient âgés de 1 an à 2 ans 11 mois, 8 enfants étaient âgés de 3 ans à 4 ans 11 mois, et 3 enfants étaient inclus dans la dernière tranche d'âge.

4. Intérêts de l'étude

Les apports de notre étude sont réels pour la création d'un outil de dépistage des TAP, car nous avons pu achever l'étalonnage du QUINOA, amorcé par Gendreau en 2024. Nous avons également développé un modèle de régression logistique, qui permet de traiter les données recueillies en minimisant le risque d'erreur. De plus, cela nous a permis de calculer les métriques de cet outil de décision, qui ont mis en évidence son caractère fiable.

Grâce à la régularisation, certaines questions qui n'étaient pas considérées comme pertinentes au vu des résultats de l'étude ont pu être extraites. Cela contribuera à réduire le temps de passation. En effet, le temps moyen de passation a été de 8 minutes et 3 secondes. Un temps de passation plus court permettrait aux médecins d'y avoir recours plus régulièrement. Ainsi, plus le QUINOA sera court et ciblé, plus il sera utilisé en pratique clinique. Nous avons choisi un coefficient de régularisation de 0,7 car cette valeur permet de ne pas modifier les métriques de performances du modèle de régression logistique, tout en supprimant 13 questions. Cela est un choix de notre part, pour que la passation du QUINOA soit nettement plus rapide, mais cela peut présenter un biais du fait de l'échantillon restreint.

5. Perspectives du travail autour du QUINOA

Actuellement, personne à notre connaissance ne s'est volontaire pour poursuivre le travail autour du QUINOA.

Nous allons tout de même détailler les perspectives de travail qui restent à accomplir dans le but que le QUINOA soit publié un jour.

5.1. Poursuite de l'étalonnage du QUINOA

En arrivant au terme de l'étude, nous avons constaté que le biais principal de cette dernière était la faible quantité de réponses reçues. Ainsi, il semblerait intéressant de poursuivre l'étalonnage du QUINOA, auprès d'enfants âgés de 1 an à 6 ans, porteurs d'un TAP ainsi que d'autres ne présentant pas ce trouble. Pour rappel, nous avons obtenu 20 réponses exploitables, et Gendreau (2024) en avait obtenu 41. Or, nous savons que, plus il y aura de participants aux phases d'étalonnage, plus la comparaison à la norme sera fiable. Ainsi, il semble primordial que l'étalonnage du QUINOA soit poursuivi pour assurer une réelle fiabilité aux professionnels de santé l'utilisant. La régression logistique permet de poursuivre l'étalonnage, tout en conservant les données recueillies jusqu'à présent. Nous rappelons que ce modèle d'apprentissage statistique se nourrit des nouveaux profils et adapte continuellement ses résultats.

De plus, il semble nécessaire de créer une page internet ou un logiciel sur lequel les professionnels de santé pourront déposer les réponses de leur patient, et auront immédiatement la décision de l'outil. Ce logiciel ou cette page internet abritera la régression logistique, qui permettra de donner la décision finale.

5.2. Poursuite du recueil des mesures psychométriques du QUINOA

Comme nous l'avons vu précédemment, deux aspects psychométriques n'ont pas pu être évalués pour le QUINOA. Ainsi, il semblerait pertinent de poursuivre ce travail sur le QUINOA en calculant la fiabilité test-retest et la fidélité inter-juges. Cela apporterait des précisions quant à la fiabilité de l'outil.

Conclusion

Nous arrivons au terme de notre étude, qui avait pour objectif d'effectuer la deuxième partie de l'étalonnage du QUINOA, concernant des enfants âgés de 1 an à 6 ans porteurs d'un TAP. Pour ce faire, nous avons créé un modèle de régression logistique, qui permet de détecter rapidement et avec fiabilité si un enfant est porteur d'un TAP ou non. Cet outil a été testé comme très sensible et très spécifique.

Cependant, face aux difficultés de recrutement auxquelles nous avons fait face, l'outil n'a pu prouver sa fiabilité, car le faible nombre de participants ayant répondu à l'étude entraîne un biais méthodologique. Ainsi, il semblerait pertinent de poursuivre l'étalonnage du QUINOA, afin que la population d'étalonnage soit plus importante, et de ce fait plus représentative des populations étudiées.

Notre étude a également permis d'écourter le temps de passation du QUINOA grâce à la régularisation I1, qui a ciblé les questions les moins pertinentes au dépistage d'un TAP. Nous avons fait le choix de réduire le QUINOA de 28 questions à 15 questions. Cela permettra une plus grande utilisation en pratique lorsqu'il sera publié.

Ce travail d'étalonnage a contribué à l'élaboration d'un test de dépistage des TAP chez les enfants porteurs d'un TAP, âgés de 1 an à 6 ans. Ce questionnaire a pour but de dépister le plus précocement possible un TAP, car ce trouble est généralement diagnostiqué tardivement. Grâce à un dépistage précoce, une prise en soin pourra être proposée dès le plus jeune âge de l'enfant par un orthophoniste, ce qui contribuera à limiter les répercussions du trouble.

Personnellement, cette étude nous a permis d'enrichir nos connaissances au sujet des TAP et des méthodes d'étalonnages, et de nous initier au monde de la recherche. Nous souhaiterions que le projet de mise au point du QUINOA soit poursuivi par un étudiant en orthophonie afin qu'il soit amélioré et enfin publié et reconnu comme un outil valide de dépistage des TAP.

Références bibliographiques

- Abadie, V. (2004). L'approche diagnostique face à un trouble de l'oralité du jeune enfant. *Archives de Pédiatrie*, 11(6), 603-605. <https://doi.org/10.1016/j.arcped.2004.03.040>
- Bardousse, L. (2019). Troubles de l'oralité alimentaire : Poursuite de la création d'un outil de dépistage à destination des médecins [Mémoire d'orthophonie]. Université de Limoges, ILFOMER.
- Barral, L. (2018). Les troubles de l'oralité : Vers la fin du dépistage tardif ? [Mémoire d'orthophonie]. Université de Limoges, ILFOMER.
- Bellis, F., Buchs-Renner, I., Vernet, M. (2009). De l'oralité heureuse à l'oralité difficile. Prévention et prise en charge dans un pôle de pédiatrie. *Spirale*, 51(3), 55-61.
<https://doi.org/10.3917/spi.051.0055>
- Comité consultatif national d'éthique. (19 mai 2019). Données massives et santé : une nouvelle approche des enjeux éthiques. https://www.ouvrirlascience.fr/wp-content/uploads/2019/06/CCNE_Donnees-massives-et-sant%C3%A9_avis130_29mai2019.pdf
- Corbière, M., & Larivière, N. (2020). *Méthodes qualitatives, quantitatives et mixtes*, 2e édition : *Dans la recherche en sciences humaines, sociales et de la santé* (Presses de l'Université du Québec).
- Demonteil, H. (2022). *Poursuite de la création d'un outil de dépistage des troubles alimentaires pédiatriques : Pré-test de la version consensuelle du QUINOA élaborée par Grueau* (2020) [Mémoire d'orthophonie]. Université de Limoges, ILFOMER.
- DeVellis, R. F. (2019). *Scale Development : Theory and Applications* (4^e éd). Sage Publications.

Dubedout, S., Cascales, T. (2016). Troubles du comportement alimentaire restrictifs du nourrisson et du jeune enfant : Situations à risque et facteurs favorisants. *Archives de Pédiatrie*, 570-576.

Farebrother, R. W. (2022). Notes on the prehistory of principal components analysis. *Journal of Multivariate Analysis*, 188, 104814. <https://doi.org/10.1016/j.jmva.2021.104814>

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.

Gedda, M. (2015). Traduction française des lignes directrices STARD pour l'écriture et la lecture des études sur la précision des tests diagnostiques. *Kinesither Revue*, 15(157), 45-49.
<https://doi.org/10.1016/j.kine.2014.11.006>

Gelas, F., & Monier-Guillaumin, H. (2016). Crédit d'un outil de dépistage des troubles de l'oralité alimentaire : Grille parentale exploratrice pour des enfants âgés de 24 à 36 mois [Mémoire d'orthophonie]. Université Claude Bernard.

Gendreau, E. (2024). *Poursuite de la création d'un outil de dépistage des troubles alimentaires pédiatriques : Construction d'un scoring (étalonnage) de la version consensuelle du Questionnaire d'investigation de l'oralité alimentaire (QUINOA)*. [Mémoire d'orthophonie]. Université de Limoges, ILFOMER.

Goday, P., Huh, S., Silverman, A., Lukens, C., Dodrill, P., Cohen, S., Delaney, A., Feuling, M., Noel, R., Gisel, E., Kenzer, A., Kessler, D., Kraus de Camargo, O., Browne, J., Phalen, J. (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>

Goulet, O., Vidailhet, M., Turck, D. (2012). *Alimentation de l'enfant en situations normale et pathologique*. Doin.

Grevesse, P., Winghem, J. V., Franck, L., Dassy, M., Cormann, N., Charlier, D., & Hermans, D. (2020). Le trouble alimentaire pédiatrique. *Percentile*, 25(2).

Grueau, P. (2020). Evaluation et révision du questionnaire de Barral (2018) en vue de produire un outil de dépistage des troubles de l'oralité alimentaire : Le QUINOA (Questionnaire d'investigation de l'oralité alimentaire) [Mémoire d'orthophonie]. Université de Limoges, ILFOMER.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: Springer.

Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. 25(1). pp. 1-18.

Lecoufle, A. (2020). Le bilan orthophonique des fonctions alimentaires du nourrisson (0-6 mois). Rééducation orthophonique, 281, 7-32.

Le Heuzey, M.-F. (2006). La prévention des troubles du comportement alimentaire du jeune enfant est-elle possible ? *Journal de Pédiatrie et de Puériculture*, 19(7), 261-264.
<https://doi.org/10.1016/j.jpp.2006.08.006>

Loubert, A., Regnault, A., Sebille, V., Cano, S., & Hardouin, J. (2020). Doit-on étalonner les instruments PRO dans les essais cliniques ? Résultat d'une étude de simulation dans le cadre de la théorie de la mesure de Rasch. *Revue d'Épidémiologie et de Santé Publique*, 68, S103-S104.

Martigne, L., Delaage, P. H., Thomas-Delecourt, F., Barthélémy, P., & Gottrand, F. (2009). P. 43 Prévalence du reflux gastro-œsophagien (RGO) chez l'enfant et l'adolescent en France: résultats d'une étude observationnelle transversale. *Gastroentérologie Clinique et Biologique*, 33(3), A40.

Mottrie, C., Moureau, A., Duret, I., Delvenne, V. (2022). Troubles des conduites alimentaires et périnatalité. *Périnatalité*, 14(1), 34-40. <https://doi.org/10.3166/rmp-2022-0148>

Quentin F. (2015). Les troubles de l'oralité chez l'ancien prématuré. *Contraste*, 41(1), 271-280.

Ramsay, M., Martel, C., Porporino, M., & Zygmuntowicz, C. (2011). The Montreal Children's Hospital Feeding Scale : A brief bilingual screening tool for identifying feeding problems. *Paediatrics & Child Health*, 16(3), 147-e17.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3077303/>

Rofidal, T. (2022). Oralité et développement de l'enfant. Dans *L'alimentation de la personne polyhandicapée* (p. 19 à 26). Érès

Saporta, G., & Niang, N. (2003). Analyse en composantes principales. *Analyse des données*, 19-42.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288.

Thibault, C. (2015). L'oralité positive. *Dialogue*, 209(3), 35-48.

<https://doi.org/10.3917/dia.209.0035>

Thibault, C. (2017). *Orthophonie et oralité: la sphère oro-faciale de l'enfant*. Elsevier Health Sciences.

Torchin, H., Ancel, P. Y., Jarreau, P. H., & Goffinet, F. (2015). Épidémiologie de la prématurité: prévalence, évolution, devenir des enfants. *Journal de gynécologie obstétrique et biologie de la reproduction*, 44(8), 723-731.

Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques : Implications pour la recherche en langue française. *Canadian Psychology / Psychologie Canadienne*, 30(4), 662-689.

Sitographie

Hôpital Necker-Enfants Malades (s. d). *Investigation des difficultés alimentaires du jeune enfant :*

Validation d'un questionnaire original : ORALQUEST. APHP. Consulté le 14 avril 2025, à l'adresse <https://www.aphp.fr/registre-des-essais-cliniques/investigation-des-difficultes-alimentaires-du-jeune-enfant-validation>

Annexes

Annexe I. Formulaire de consentement à destination des parents	55
Annexe II. Version consensuelle du QUINOA	57
Annexe III. Nouvelles questions ajoutées au QUINOA.....	61
Annexe IV. Nouvelle version du QUINOA	62
Annexe V. Affiche de présentation du QUINOA	67
Annexe VI. Mails de diffusion du QUINOA envoyés aux professionnels de santé	68
Annexe VI.I. Novembre 2024 : Premier contact.....	68
Annexe VI.II. Décembre 2024 : Première relance.....	69
Annexe VI.III. Janvier 2025 : Deuxième relance	70
Annexe VII. Répartition des scores selon Gendreau (2024).....	71
Annexe VIII. Script prétraitement	76
Annexe IX. Script d'apprentissage de la régression logistique et de la régularisation l1	82
Annexe X. Script ACP	83
Annexe XI. Script courbes ROC.....	84
Annexe XII. Graphiques des résultats obtenus à chaque question du QUINOA.....	85
Annexe XIII. Poids des questions de la régression logistique.....	95
Annexe XIV. Poids des questions de la régression logistique avec la régularisation l1	96

Annexe I. Formulaire de consentement à destination des parents



Etalonnage du QUINOA

ILFOMER Institut des sciences de la réadaptation

39F Rue CAMILLE GUERIN
87036 LIMOGES

Investigateur

Treyt Chiara
chiara.treyt@etu.unilim.fr

Mémoire dirigé par Olivier Prot, maître de conférences

NOTE D'INFORMATION AUX PERSONNES

Madame, Monsieur,

Nous proposons à votre enfant de participer à une étude dont l'objectif est détailler un outil de dépistage des troubles alimentaires pédiatriques : le QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire). Ce travail rentre dans le cadre d'un mémoire d'orthophonie.

Dans notre étude, nous avons pour objectif de recueillir de nombreuses données, qui nous permettront de déterminer des seuils d'alerte, les plus fiables possibles, au dépistage des troubles alimentaires pédiatriques.

Si vous acceptez de participer à notre étude, il vous sera demandé de répondre à quelques questions concernant l'oralité de votre enfant, et notamment son rapport à l'alimentation. Cette expérimentation devrait inclure environ 50 enfants âgés de 1 an à 6 ans, en France.

Pourquoi cette étude ?

Les troubles alimentaires pédiatriques peuvent avoir des conséquences sur le développement physique et la santé des enfants. Ainsi, un dépistage précoce permettrait de réduire ces conséquences. L'objectif du QUINOA est que les professionnels de santé puissent dépister au mieux un trouble alimentaire pédiatrique afin de pouvoir orienter l'enfant et sa famille le plus rapidement possible chez un(e) orthophoniste pour rééduquer ce trouble et ainsi de limiter les conséquences sur le développement de l'enfant.

Comment va se dérouler cette étude ?

Nous recrutons des enfants de 1 an à 6 ans, dont le diagnostic de trouble alimentaire pédiatrique a été posé. Nous allons administrer le QUINOA aux parents de ces enfants. Cela se fera par le biais d'un professionnel de santé. Le rôle des parents sera de répondre à des questions concernant leur enfant. Le questionnaire est composé de 22 questions au total. Une analyse quantitative des données recueillies sera ensuite réalisée par l'investigatrice. Cette étude ne comporte aucun risque prévisible pour la santé.

Quels sont vos droits ?

Votre participation à cette étude est entièrement libre. Vous pouvez ne pas prendre part à cette étude ou si vous souhaitez vous en retirer à quelque moment que ce soit, et quel que soit le motif.

Les données recueillies au cours de cette étude resteront strictement confidentielles, vos noms et adresse ne seront pas divulgués. Compte tenu des nécessités de la recherche et de son analyse ultérieure, les données recueillies qui concernent votre enfant feront l'objet d'un traitement informatisé et anonyme. Les informations qui feront l'objet de ce traitement informatisé sont les suivantes : les informations remplies au préalable par le professionnel de santé concernant le patient (date de naissance, taille, poids, tonicité de la sphère oro-faciale) et les réponses des parents au test. L'article 40 de la loi prévoit votre droit d'accès, d'opposition et de rectification des données enregistrées sur informatique, à tout moment, par l'intermédiaire du responsable de l'étude : Chiara Treyt, 15 rue des Tanneries 87000 Limoges. 06.52.09.35.28.

Lorsque cette étude sera terminée, si vous le souhaitez, vous serez personnellement informé des résultats des examens et de l'étude par l'investigatrice, dès que ceux-ci seront disponibles. Après avoir lu cette note d'information, n'hésitez pas à poser toutes les questions que vous désirez.

Si vous acceptez que votre enfant participe à cette étude, il vous suffit, comme le prévoit la loi, de signer le consentement de participation prévu à cet effet, en triple exemplaire. Un exemplaire vous sera remis.

Nous vous remercions de votre collaboration.

ACCORD DE PARTICIPATION

Je soussigné(e).....

Né(e) le/...../.....

Déclare :

- qu'il a été proposé à mon enfant de participer à l'étude sus-nommée,
- que m'a été communiqué : l'objectif, la méthode, et la durée de l'étude, mon droit de refuser de participer ou de retirer mon consentement à tout moment sans encourir aucune responsabilité,
- que le protocole m'a été expliqué en détail,
- que j'ai répondu en toute bonne foi aux questions concernant l'état de santé de mon enfant et sa participation à d'autres études.

Après avoir discuté librement et obtenu réponses à toutes mes questions, j'accepte, en toute connaissance de cause, que mon enfant participe à l'étude.

Fait à....., le.....

Nom du/des parent(s) :

Nom de l'expérimentateur :

Signature du/des parent(s)
(précédée de la mention « lu et approuvé ») :

Signature de l'expérimentateur :

Annexe II. Version consensuelle du QUINOA

Partie réservée au médecin

Date de naissance de l'enfant :

jj/mm/aaaa 

Poids :



Taille (en cm) :



1. Lèvres/langue/joues/mâchoire de l'enfant sont-elles toniques ?

- Pas ou peu toniques : bavage Tonus normal Hypertoniques : tête projetée en haut et mâchoire Ne sait pas exagérément ouverte

Habiletés oro-motrices

2. L'enfant a-t-il (ou a-t-il eu) :

Plusieurs réponses possibles.

- Des difficultés à téter Des difficultés à mâcher Aucune de ces difficultés

3. L'enfant a-t-il (ou a-t-il eu) besoin d'interrompre la prise alimentaire parce qu'il a (ou avait) l'air fatigué ou en difficulté ?

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Ne sait pas

4. L'enfant accumule-t-il de la nourriture dans ses joues et/ou la recrache après l'avoir stockée ?

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

5. En général, le repas de l'enfant dure en moyenne :

- Moins de 15 minutes Entre 15 et 30 minutes Entre 30 et 45 minutes Entre 45 minutes et une heure Plus d'une heure

Sensorialité

6. Estimez-vous que votre enfant est plus "casse-cou" que les autres enfants de son âge ?

Par exemple : marche sur les jouets sans s'en rendre compte, aime se balancer, aime les jeux qui tournent comme les manèges, ne tient pas en place sur sa chaise, grimpe partout ... comportements pouvant aller jusqu'à une prise de risques, une mise en danger

- Beaucoup plus "casse-cou" Un peu plus "casse-cou" Pareil que les autres enfants Plus prudent Ne sait pas

7. A quelle fréquence l'enfant :

- refuse qu'on lui touche le visage ou le corps (caresses, bain, câlins) ou
- refuse le contact des mains, des pieds ou du corps avec certaines matières (aliments, textiles, herbe, sable, pâte à modeler...)

- Jamais Rarement Parfois, dans certaines situations Assez souvent Très souvent

8. L'enfant porte-t-il (ou a-t-il porté) les objets ou ses doigts à la bouche en dehors des périodes de poussées dentaires ?

- Exceptionnellement : l'enfant met (ou a mis) rarement en bouche Parfois : pour s'endormir ou se rassurer Souvent Très fréquemment et tout type d'objets

9. Lors du repas ou de l'approche d'un objet ou aliment (en dehors des aliments favoris ou acceptés depuis toujours) de la bouche, à quelle fréquence l'enfant manifeste-t-il l'un ou plusieurs de ces comportements ?

Par exemple : Grimaces de dégoût, réflexes nauséens importants voire vomissements, frissons, réactions de retrait, pleurs, agitation, recrachement de la nourriture.

- Systématiquement 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

10. A quelle fréquence l'enfant a-t-il des haut-le-cœur lors de l'introduction d'un aliment (en dehors des aliments favoris ou acceptés depuis toujours) ou d'un objet en particulier dans sa bouche ?

- Systématiquement 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

L'alimentation de l'enfant

11. Diriez-vous que l'introduction de nouvelles textures (passage aux purées ou passage aux morceaux) a été :

- Facile Ni facile ni difficile Assez difficile Très difficile Impossible

12. L'enfant refuse-t-il de goûter des aliments nouveaux, choisit les aliments qu'il peut manger et trie son assiette ?

Par exemple : selon leur texture (purée, morceaux...) / leur couleur / leur odeur / leur forme / leur température (froid, tiède, chaud) / leur marque / leur emballage / la personne qui a cuisiné...

- Jamais Rarement Parfois Assez souvent Très souvent

13. L'enfant a-t-il une préférence pour les aliments au goût prononcé ? (Par exemple : cornichons, moutarde ...)

ou

L'enfant prend-il une grande quantité d'aliments en bouche ?

- Jamais Rarement Parfois Assez souvent Très souvent

Le comportement de l'enfant et de sa famille lors des repas

14. L'enfant manifeste-t-il sa faim ?

- Jamais Rarement Parfois Assez souvent Très souvent

15. L'enfant redoute-t-il les moments de repas, le passage à table se fait-il difficilement ?

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

16. L'enfant a-t-il des attitudes de défense, d'évitement envers les aliments (en dehors des aliments favoris ou acceptés depuis toujours) ?

Par exemple : ferme les lèvres pour ne pas laisser entrer la cuillère, détourne la tête et/ou le corps, se protège le visage avec la main, repousse la cuillère, refuse de manger, transpire, s'agit.

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

17. L'enfant réagit-il différemment (refus de manger...) selon la personne avec qui il mange ?

- Jamais Rarement Parfois Assez souvent Très souvent

18. L'enfant a-t-il peur d'avaler et/ou de s'étouffer ?

- A chaque repas Plusieurs fois par semaine Moins d'une fois par semaine Jamais Ne sait pas

19. Les repas sont-ils difficiles, sources de conflits et d'anxiété pour vous et votre enfant ?

- Jamais Rarement Parfois Assez souvent Très souvent

20. Utilisez-vous des stratégies dans le but que votre enfant mange ?

Par exemple : utiliser la ruse, lui faire du chantage, le distraire (télévision, jouets...), le forcer à goûter, le cajoler, modifier son menu...

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

Etat de santé

21. L'enfant fait-il des fausses-routes (la nourriture passe de travers) ?

Par exemple : il tousse, semble s'étouffer.

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

22. Enrichissez-vous l'alimentation de votre enfant parce que vous avez l'impression qu'il ne mange pas assez ?

Par exemple : Ajouter des matières grasses, du sucre... ou des compléments alimentaires.

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

Annexe III. Nouvelles questions ajoutées au QUINOA

QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire)

Prise en soin / Rééducation

L'enfant bénéficie-t-il d'une rééducation (orthophonique, kinésithérapie, etc.) actuellement ?

Oui

Non

De quelle rééducation s'agit-il ?

Date de début de la rééducation :

jj/mm/aaaa

Annexe IV. Nouvelle version du QUINOA

QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire)

Prise en soin / Rééducation

1. L'enfant bénéficie-t-il d'une rééducation (orthophonique, kinésithérapie, etc.) actuellement ?

Oui Non

1bis. De quelle rééducation s'agit-il ?

1ter. Date de début de la rééducation :

 jj/mm/aaaa 

Partie réservée au médecin

2. Date de naissance de l'enfant :

 jj/mm/aaaa 

3. Poids :

 kg 

4. Taille (en cm) :

 cm 

5. Lèvres/langue/joues/mâchoire de l'enfant sont-elles toniques ?

- Pas ou peu toniques : bavage Tonus normal Hypertoniques : tête projetée en haut et mâchoire Ne sait pas
exagérément ouverte

Habiletés oro-motrices

6. L'enfant a-t-il (ou a-t-il eu) :

Plusieurs réponses possibles.

- Des difficultés à téter Des difficultés à mâcher Aucune de ces difficultés

7. L'enfant a-t-il (ou a-t-il eu) besoin d'interrompre la prise alimentaire parce qu'il a (ou avait) l'air fatigué ou en difficulté ?

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Ne sait pas

8. L'enfant accumule-t-il de la nourriture dans ses joues et/ou la recrache après l'avoir stockée ?

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

9. En général, le repas de l'enfant dure en moyenne :

- Moins de 15 minutes Entre 15 et 30 minutes Entre 30 et 45 minutes Entre 45 minutes et une heure Plus d'une heure

Sensorialité

10. Estimez-vous que votre enfant est plus "casse-cou" que les autres enfants de son âge ?

Par exemple : marche sur les jouets sans s'en rendre compte, aime se balancer, aime les jeux qui tournent comme les manèges, ne tient pas en place sur sa chaise, grimpe partout ... comportements pouvant aller jusqu'à une prise de risques, une mise en danger

- Beaucoup plus "casse-cou" Un peu plus "casse-cou" Pareil que les autres enfants Plus prudent Ne sait pas

11. A quelle fréquence l'enfant :

- refuse qu'on lui touche le visage ou le corps (caresses, bain, câlins) ou
- refuse le contact des mains, des pieds ou du corps avec certaines matières (aliments, textiles, herbe, sable, pâte à modeler...)

- Jamais Rarement Parfois, dans certaines situations Assez souvent Très souvent

12. L'enfant porte-t-il (ou a-t-il porté) les objets ou ses doigts à la bouche en dehors des périodes de poussées dentaires ?

- Exceptionnellement : l'enfant met (ou a mis) rarement en bouche Parfois : pour s'endormir ou se rassurer Souvent Très fréquemment et tout type d'objets

13. Lors du repas ou de l'approche d'un objet ou aliment (en dehors des aliments favoris ou acceptés depuis toujours) de la bouche, à quelle fréquence l'enfant manifeste-t-il l'un ou plusieurs de ces comportements ?

Par exemple : Grimaces de dégoût, réflexes nauséens importants voire vomissements, frissons, réactions de retrait, pleurs, agitation, recrachement de la nourriture.

- Systématiquement 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

14. A quelle fréquence l'enfant a-t-il des haut-le-cœur lors de l'introduction d'un aliment (en dehors des aliments favoris ou acceptés depuis toujours) ou d'un objet en particulier dans sa bouche ?

- Systématiquement 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

L'alimentation de l'enfant

15. Diriez-vous que l'introduction de nouvelles textures (passage aux purées ou passage aux morceaux) a été :

- Facile Ni facile ni difficile Assez difficile Très difficile Impossible

16. L'enfant refuse-t-il de goûter des aliments nouveaux, choisit les aliments qu'il peut manger et trie son assiette ?

Par exemple : selon leur texture (purée, morceaux...) / leur couleur / leur odeur / leur forme / leur température (froid, tiède, chaud) / leur marque / leur emballage / la personne qui a cuisiné...

- Jamais Rarement Parfois Assez souvent Très souvent

17. L'enfant a-t-il une préférence pour les aliments au goût prononcé ? (Par exemple : cornichons, moutarde ...)

ou

L'enfant prend-il une grande quantité d'aliments en bouche ?

- Jamais Rarement Parfois Assez souvent Très souvent

Le comportement de l'enfant et de sa famille lors des repas

18. L'enfant manifeste-t-il sa faim ?

Jamais Rarement Parfois Assez souvent Très souvent

19. L'enfant redoute-t-il les moments de repas, le passage à table se fait-il difficilement ?

A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

20. L'enfant a-t-il des attitudes de défense, d'évitement envers les aliments (en dehors des aliments favoris ou acceptés depuis toujours) ?

Par exemple : ferme les lèvres pour ne pas laisser entrer la cuillère, détourne la tête et/ou le corps, se protège le visage avec la main, repousse la cuillère, refuse de manger, transpire, s'agitte.

A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

21. L'enfant réagit-il différemment (refus de manger...) selon la personne avec qui il mange ?

Jamais Rarement Parfois Assez souvent Très souvent

22. L'enfant a-t-il peur d'avaler et/ou de s'étouffer ?

A chaque repas Plusieurs fois par semaine Moins d'une fois par semaine Jamais Ne sait pas

23. Les repas sont-ils difficiles, sources de conflits et d'anxiété pour vous et votre enfant ?

Jamais Rarement Parfois Assez souvent Très souvent

24. Utilisez-vous des stratégies dans le but que votre enfant mange ?

Par exemple : utiliser la ruse, lui faire du chantage, le distraire (télévision, jouets...), le forcer à goûter, le cajoler, modifier son menu...

A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

Etat de santé**25. L'enfant fait-il des fausses-routes (la nourriture passe de travers) ?**

Par exemple : il tousse, semble s'étouffer.

A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

26. Enrichissez-vous l'alimentation de votre enfant parce que vous avez l'impression qu'il ne mange pas assez ?

Par exemple : Ajouter des matières grasses, du sucre... ou des compléments alimentaires.

- A chaque repas 1 fois par jour Plusieurs fois par semaine Moins d'une fois par semaine Jamais

Annexe V. Affiche de présentation du QUINOA

Chiara Treyt
Etudiante en 5ème année d'orthophonie à Limoges
chiara.treyt@etu.unilim.fr
06.52.09.35.28

Mémoire dirigé par :
M. Olivier Prot
Maître de conférence

QUINOA

Questaionnaire d'INvestigation de l'Oralité Alimentaire

Les troubles alimentaires pédiatriques

Les Troubles Alimentaire Pédiatriques (TAP) représentent l'ensemble des difficultés à s'alimenter par la bouche. Les TAP peuvent se manifester de différentes manières : désintérêt à s'alimenter, hypo ou hyper sensibilité aux odeurs, couleurs, goûts, etc. Ces manifestations peuvent avoir de lourdes conséquences concernant le développement physique de l'enfant, son état de santé et/ou ses interactions sociales.

Une prise en soin précoce permet de limiter l'impact des troubles et leur chronicisation, ce qui explique le besoin de créer un outil permettant de dépister ce trouble, le plus précocement possible. Le QUINOA a été créé dans le but de palier ce manque.



Le QUINOA

Louise Barral (2018), Léa Bardousse (2019), Pauline Grueau (2020), Hélène Demonteil (2022) et Eléonore Gendreau (2024) se sont penchées sur la conception du QUINOA et ont effectué les premières étapes de la conception de ce questionnaire. Aujourd'hui, l'objectif de notre travail est d'effectuer la deuxième partie de l'étalonnage de cet outil, c'est à dire d'administrer le QUINOA à des enfants âgés de 1 à 6 ans, porteurs d'un TAP. Nous avons choisi de scinder le QUINOA en 3 tranches d'âge : 1 an à 2 ans 11 mois, 3 ans à 4 ans 11 mois et 5 ans à 6 ans. Cela permettra de mettre en avant des signes d'alerte représentatifs pour chaque classe d'âge.

Ainsi, nous vous sollicitons afin de pouvoir recueillir le plus de résultats possibles, ce qui nous permettra de créer un outil fiable et valide.

Nous vous remercions par avance pour votre aide !

Votre rôle :

Administrer le QUINOA au(x) parent(s) d'un ou plusieurs patient(s) âgé(s) de 1 an à 6 ans avec un diagnostic de trouble alimentaire pédiatrique.



Les professionnels de santé sollicités :

- Médecins généralistes
- Pédiatres
- Médecins ou pédiatres en PMI
- Médecins scolaires
- Dentistes
- Orthodontistes
- Diététiciens
- Kinésithérapeutes
- Ergothérapeutes
- Psychomotriciens
- Orthophonistes
- Psychologues
- Psychiatres
- Pédopsychiatres
- Gastro-pédiatres
- Gastro-entérologues
- Oto-rhino-laryngologues

Annexe VI. Mails de diffusion du QUINOA envoyés aux professionnels de santé

Annexe VI.I. Novembre 2024 : Premier contact

Madame, Monsieur,

Je suis étudiante en 5^{ème} année d'orthophonie à Limoges. Mon mémoire de fin d'étude porte sur la poursuite de la création d'un questionnaire de dépistage des troubles alimentaires pédiatriques : le **QUINOA** (QuesTionnaire d'INvestigation de l'Oralité Alimentaire). Vous trouverez en pièce jointe une présentation de ce projet.

Ce travail s'inscrit dans la continuité des travaux de Louise Barral, Léa Bardousse, Pauline Grueau, Hélène Demonteil et Eléonore Gendreau, auxquels certains d'entre vous ont peut-être déjà participé.

Ce travail consiste en l'**élaboration de la deuxième partie de l'étalonnage** du QUINOA : le recueil des données chez des sujets porteurs d'un trouble alimentaire pédiatrique. Ainsi, je vous sollicite aujourd'hui pour vous proposer de participer à l'élaboration de cet étalonnage. Si vous acceptez de nous aider dans ce travail, votre objectif sera d'administrer le QUINOA au(x) parent(s) d'un ou plusieurs patient(s) âgé(s) de 1 an à 6 ans avec un **diagnostic de trouble alimentaire pédiatrique posé**. Vous devrez également leur faire remplir le formulaire de consentement, avant la passation. Le questionnaire est anonyme.

Votre aide dans notre étude est précieuse, car la récolte de ces données permettra de déterminer un seuil d'alerte permettant de dépister un risque de trouble alimentaire pédiatrique chez un enfant tout-venant. Ainsi, chaque questionnaire rempli fera grandement avancer notre travail.

Pour accéder au questionnaire, plusieurs options s'offrent à vous :

- Format numérique, via ce lien : <https://sphinx.unilim.fr/v4/s/7uoshf>;
- Format PDF pour impression sur simple demande de votre part ;
- Version papier, par voie postale avec une enveloppe timbrée pour le retour, sur simple demande de votre part.

Si vous souhaitez participer à cette étude, merci de nous confirmer votre participation en répondant à ce mail. Ce recueil de données est ouvert jusqu'au 31 janvier 2025.

Si vous connaissez d'autres professionnels de santé susceptibles d'être intéressés par ce projet, n'hésitez pas à leur transférer ce mail.

Je reste à votre disposition pour toute information complémentaire.

Merci d'avance pour l'intérêt que vous porterez à ce projet et le temps que vous y consacrerez.

Bien cordialement,

Chiara Treyt,
Etudiante en 5^{ème} année d'orthophonie au CFUO de Limoges
Chiara.treyt@etu.unilim.fr

Annexe VI.II. Décembre 2024 : Première relance

Madame, Monsieur,

Je suis étudiante en 5^{ème} année d'orthophonie à Limoges. Mon mémoire de fin d'étude porte sur la poursuite de la création d'un questionnaire de dépistage des troubles alimentaires pédiatriques : le **QUINOA** (Q**U**estionnaire d'**I**Nvestigation de l'**O**ralité **A**limentaire). Vous trouverez en pièce jointe une présentation de ce projet.

Vous avez peut-être déjà reçu une sollicitation de ma part au mois de novembre 2023. Je me permets de vous solliciter à nouveau car nous avons reçu très peu de réponses, et nous savons que plus nous aurons de réponses, plus l'outil de dépistage en cours de création sera fiable. Ainsi, chaque participation est importante et nous aidera grandement dans notre étude.

Ce travail consiste en **l'élaboration de la deuxième partie de l'étalement du QUINOA** : le recueil des données chez des sujets porteurs d'un trouble alimentaire pédiatrique. Ainsi, je vous sollicite aujourd'hui pour vous proposer de participer à l'élaboration de cet étalement. Si vous acceptez de nous aider dans ce travail, votre objectif sera d'administrer le QUINOA au(x) parent(s) d'un ou plusieurs patient(s) âgé(s) de 1 an à 6 ans avec un **diagnostic de trouble alimentaire pédiatrique posé**. Vous devrez également leur faire remplir le formulaire de consentement, avant la passation. Le questionnaire est anonyme.

Votre aide dans notre étude est précieuse, car la récolte de ces données permettra de déterminer un seuil d'alerte permettant de dépister un risque de trouble alimentaire pédiatrique chez un enfant tout-venant. Ainsi, chaque questionnaire rempli fera grandement avancer notre travail.

Pour accéder au questionnaire, plusieurs options s'offrent à vous :

- Format numérique, via ce lien : <https://sphinx.unilim.fr/v4/s/7uoshf>;
- Format PDF pour impression sur simple demande de votre part ;
- Version papier, par voie postale avec une enveloppe timbrée pour le retour, sur simple demande de votre part.

Si vous souhaitez participer à cette étude, merci de nous confirmer votre participation en répondant à ce mail. Ce recueil de données est ouvert jusqu'au 31 janvier 2025.

Si vous connaissez d'autres professionnels de santé susceptibles d'être intéressés par ce projet, n'hésitez pas à leur transférer ce mail.

Je reste à votre disposition pour toute information complémentaire.

Merci d'avance pour l'intérêt que vous porterez à ce projet et le temps que vous y consacrerez.

Bien cordialement,

Chiara Treyt,
Etudiante en 5^{ème} année d'orthophonie au CFUO de Limoges
Chiara.treyt@etu.unilim.fr

Annexe VI.III. Janvier 2025 : Deuxième relance

Madame, Monsieur,

Je suis étudiante en 5ème année d'orthophonie à Limoges. Mon mémoire de fin d'étude porte sur la poursuite de la création d'un questionnaire de dépistage des troubles alimentaires pédiatriques : le QUINOA (~~Q~~uestionnaire d'~~I~~nvestigation de l'Oralité Alimentaire). Vous trouverez en pièce jointe une présentation de ce projet.

Vous avez peut-être déjà reçu une sollicitation de ma part aux mois de novembre et de décembre 2024. Je me permets de vous solliciter à nouveau car nous avons reçu très peu de réponses, et nous savons que plus nous aurons de réponses, plus l'outil de dépistage en cours de création sera fiable. Ainsi, chaque participation est important et nous aidera grandement dans notre étude.

Ce travail consiste en l'**élaboration de la deuxième partie de l'étalonnage du QUINOA** : le recueil des données chez des sujets porteurs d'un trouble alimentaire pédiatrique. Ainsi, je vous sollicite aujourd'hui pour vous proposer de participer à l'élaboration de cet étalonnage. Si vous acceptez de nous aider dans ce travail, votre objectif sera d'administrer le QUINOA au(x) parent(s) d'un ou plusieurs patient(s) âgé(s) de 1 an à 6 ans avec un **diagnostic de trouble alimentaire pédiatrique posé**. Vous devrez également leur faire remplir le formulaire de consentement, avant la passation. Le questionnaire est anonyme.

Votre aide dans notre étude est précieuse, car la récolte de ces données permettra de déterminer un seuil d'alerte permettant de dépister un risque de trouble alimentaire pédiatrique chez un enfant tout-venant. Ainsi, chaque questionnaire rempli fera grandement avancer notre travail.

Pour accéder au questionnaire, plusieurs options s'offrent à vous :

- Format numérique, via ce lien : <https://sphinx.unilim.fr/v4/s/7uoshf>;
- Format PDF pour impression sur simple demande de votre part ;
- Version papier, par voie postale avec une enveloppe timbrée pour le retour, sur simple demande de votre part.

Si vous souhaitez participer à cette étude, merci de nous confirmer votre participation en répondant à ce mail. Ce recueil de données est ouvert jusqu'au 31 janvier 2025.

Si vous connaissez d'autres professionnels de santé susceptibles d'être intéressés par ce projet, n'hésitez pas à leur transférer ce mail.

Je reste à votre disposition pour toute information complémentaire.

Merci d'avance pour l'intérêt que vous porterez à ce projet et le temps que vous y consacrerez.

Bien cordialement,

Chiara Treyt,
Etudiante en 5ème année d'orthophonie au CFUO de Limoges
Chiara.treyt@etu.unilim.fr

Annexe VII. Répartition des scores selon Gendreau (2024)

Tableau 4 : Répartition des points pour la question 5

Pas ou peu toniques : bavage	Tonus normal	Hypertoniques : tête projetée en haut et mâchoire exagérément ouverte	Ne sait pas	Total : /1 → /4
1 (0 pour les moins de 3 ans)	0	1	0	

Tableau 5 : Répartition des scores pour la question 6

Des difficultés à téter	Des difficultés à mâcher	Aucune de ces difficultés	Total : /2 → /4
1	1	0	

Tableau 6 : Répartition des points pour la question 7

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Ne sait pas	Total : /3 → /4
3	2	1	0	0	

Tableau 7 : Répartition des points pour la question 8

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Tableau 8 : Répartition des points pour la question 9

Moins de 15 minutes	Entre 15 et 30 minutes	Entre 30 et 45 minutes	Entre 45 minutes et une heure	Plus d'une heure	Total : /4
2	0	1	3	4	

Tableau 9 : Répartition des points pour la question 10

Beaucoup plus « casse-cou »	Un peu plus « casse-cou »	Pareil que les autres enfants	Plus prudent	Ne sait pas	Total : /3 → /4
3	1	0	2	0	

Tableau 10 : Répartition des points pour la question 11

Jamais	Rarement	Parfois, dans certaines situations	Assez souvent	Très souvent	Total : /4
0	1	2	3	4	

Tableau 11 : Répartition des points pour la question 12

Exceptionnellement : l'enfant met (ou a mis) rarement en bouche	Parfois : pour s'endormir ou se rassurer	Souvent	Très fréquemment et tout types d'objets	Total : /3 → /4
1	2	0	3	

Tableau 12 : Répartition des points pour la question 13

Systématiquement	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	0	1	

Tableau 13 : Répartition des points pour la question 14

Systématiquement	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	0	1	

Tableau 14 : Répartition des points pour la question 15

Facile	Ni facile ni difficile	Assez difficile	Très difficile	Impossible	Total : /4
0	1	2	3	4	

Tableau 15 : Répartition des points pour la question 16

Jamais	Rarement	Parfois	Assez souvent	Très souvent	Total : /4
0	1	2	3	4	

Tableau 16 : Répartition des points pour la question 17

Jamais	Rarement	Parfois	Assez souvent	Très souvent	Total : /4
0	2	1	3	4	

Tableau 17 : Répartition des points pour la question 18

Jamais	Rarement	Parfois	Assez souvent	Très souvent	Total : /3 → /4
3	2	1	1	0	

Tableau 18 : Répartition des points pour la question 19

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Tableau 19 : Répartition des points pour la question 20

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Tableau 20 : Répartition des points pour la question 21

Jamais	Rarement	Parfois	Assez souvent	Très souvent	Total : /4
0	1	2	3	4	

Tableau 21 : Répartition des points pour la question 22

A chaque repas	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Ne sait pas	Total : /3 → /4
3	2	1	0	0	

Tableau 22 : Répartition des points pour la question 23

Jamais	Rarement	Parfois	Assez souvent	Très souvent	Total : /4
0	1	2	3	4	

Tableau 23 : Répartition des points pour la question 24

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Tableau 24 : Répartition des points pour la question 25

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Tableau 25 : Répartition des points pour la question 26

A chaque repas	1 fois par jour	Plusieurs fois par semaine	Moins d'une fois par semaine	Jamais	Total : /4
4	3	2	1	0	

Annexe VIII. Script prétraitement

```
import pandas as pd
import numpy as np

# Code fait par Chiara en Préambule pour définir l'échelle utilisée
# L est une liste pour garder les infos
L = []

#####
# 0
# catégorie donnée par pd.factorize
categ =["Ne sait pas", "Tonus normal", "Pas ou peu toniques : bavage"]
# valeurs / note par catégorie
note =np.array([0, 1, 2])
# nouvelles valeurs
note_new =np.array([0, 0, 4])
print("")
print("#####")
print("Question: 1...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([0,"1...",categ,note,note_new])

#####
# 1
# catégorie donnée par pd.factorize
categ =["Aucune de ces difficultés", "Des difficultés à téter", "Des difficultés à mâcher", "Des difficultés à téter;Des difficultés à mâcher"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3])
# nouvelles valeurs
note_new =np.array([0, 2, 2, 4])
print("")
print("#####")
print("Question: 2...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([1,"2...",categ,note,note_new])

#####
# 2
# catégorie donnée par pd.factorize
categ =[ "Moins d'une fois par semaine", "Plusieurs fois par semaine", "Ne sait pas", "1 fois par jour", "A chaque repas"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1.33, 0, 2.66, 4])
print("")
print("#####")
print("Question: 3...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([2,"3...",categ,note,note_new])

#####
# 3
# catégorie donnée par pd.factorize
categ =[ "Plusieurs fois par semaine", "Jamais", "Moins d'une fois par semaine", "A chaque repas", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([2, 0, 1, 4, 3])
print("")
print("#####")
print("Question: 4...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([3,"4...",categ,note,note_new])
```

```

#####
# 4
# categorie donnée par pd.factorize
categ =["Entre 30 et 45 minutes", "Moins de 15 minutes", "Entre 15 et 30 minutes", "Entre 45 minutes et une
heure", "Plus d'une heure"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([1, 2, 0, 3, 4])
print("")
print("#####")
print("Question: 5...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([4,"5...",categ,note,note_new])

#####
# 5
# categorie donnée par pd.factorize
categ =["Beaucoup plus casse-cou", "Pareil que les autres enfants", "Un peu plus casse-cou", "Plus prudent"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3])
# nouvelles valeurs
note_new =np.array([4, 0, 1.33, 2.66])
print("")
print("#####")
print("Question: 6...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([5,"6...",categ,note,note_new])

#####
# 6
# categorie donnée par pd.factorize
categ =["Jamais", "Rarement", "Parfois, dans certaines situations", "Très souvent", "Assez souvent"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 4, 3])
print("")
print("#####")
print("Question: 7...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([6,"7...",categ,note,note_new])

#####
# 7
# categorie donnée par pd.factorize
categ =[("Souvent", "Exceptionnellement : l'enfant met (ou a mis) rarement en bouche", "Parfois : pour s'endormir
ou se rassurer", "Très fréquemment et tout type d'objets")]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3])
# nouvelles valeurs
note_new =np.array([0, 1.33, 2.66, 4])
print("")
print("#####")
print("Question: 8...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([7,"8...",categ,note,note_new])

#####
# 8
# categorie donnée par pd.factorize
categ =[("Moins d'une fois par semaine", "Jamais", "Plusieurs fois par semaine", "Systématiquement", "1 fois par
jour")]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 4, 3])
print("")
print("#####")
print("Question: 9...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([8,"9...",categ,note,note_new])

```

```

#####
# 9
# categorie donnée par pd.factorize
categ =["Moins d'une fois par semaine", "Jamais", "Plusieurs fois par semaine", "Systématiquement", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([1, 0, 2, 4, 3])
print("")
print("#####")
print("Question: 10...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([9,"10...",categ,note,note_new])

#####
# 10
# categorie donnée par pd.factorize
categ =["Ni facile ni difficile", "Facile", "Assez difficile", "Très difficile"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3])
# nouvelles valeurs
note_new =np.array([1, 0, 2, 3])
print("")
print("#####")
print("Question: 11...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([10,"11...",categ,note,note_new])

#####
# 11
# categorie donnée par pd.factorize
categ =["Rarement", "Jamais", "Parfois", "Assez souvent", "Très souvent"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 3, 4])
print("")
print("#####")
print("Question: 12...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([11,"12...",categ,note,note_new])

#####
# 12
# categorie donnée par pd.factorize
categ =["Assez souvent", "Rarement", "Très souvent", "Jamais", "Parfois"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([3, 2, 4, 0, 1])
print("")
print("#####")
print("Question: 13...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([12,"13...",categ,note,note_new])

#####
# 13
# categorie donnée par pd.factorize
categ =["Assez souvent", "Très souvent", "Parfois", "Rarement", "Jamais"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([1.33, 0, 1.33, 2.66, 4])
print("")
print("#####")
print("Question: 14...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([13,"14...",categ,note,note_new])

```

```

#####
# 14
# categorie donnée par pd.factorize
categ =["Jamais", "Moins d'une fois par semaine", "Plusieurs fois par semaine", "A chaque repas", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 4, 3])
print("")
print("#####")
print("Question: 15...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([14,"15...",categ,note,note_new])

#####
# 15
# categorie donnée par pd.factorize
categ =["Moins d'une fois par semaine", "Jamais", "Plusieurs fois par semaine", "A chaque repas", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([1, 0, 2, 4, 3])
print("")
print("#####")
print("Question: 16...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([15,"16...",categ,note,note_new])

#####
# 16
# categorie donnée par pd.factorize
categ =[ "Jamais", "Parfois", "Assez souvent", "Rarement", "Très souvent"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 2, 3, 1, 4])
print("")
print("#####")
print("Question: 17...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([16,"17...",categ,note,note_new])

#####
# 17
# categorie donnée par pd.factorize
categ =[ "Jamais", "Ne sait pas", "Moins d'une fois par semaine", "A chaque repas", "Plusieurs fois par semaine"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 0, 1.33, 4, 2.66])
print("")
print("#####")
print("Question: 18...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([17,"18...",categ,note,note_new])

#####
# 18
# categorie donnée par pd.factorize
categ =[ "Parfois", "Jamais", "Rarement", "Assez souvent", "Très souvent"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([2, 0, 1, 3, 4])
print("")
print("#####")
print("Question: 19...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([18,"19...",categ,note,note_new])

```

```

#####
# 19
# categorie donnée par pd.factorize
categ =["Jamais", "Moins d'une fois par semaine", "1 fois par jour", "Plusieurs fois par semaine", "A chaque repas"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 3, 2, 4])
print("")
print("#####")
print("Question: 20...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([19,"20...",categ,note,note_new])

#####
# 20
# categorie donnée par pd.factorize
categ =["Jamais", "Moins d'une fois par semaine", "Plusieurs fois par semaine", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 3])
print("")
print("#####")
print("Question: 21...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([20,"21...",categ,note,note_new])

#####
# 21
# categorie donnée par pd.factorize
categ =["Jamais", "Moins d'une fois par semaine", "Plusieurs fois par semaine", "A chaque repas", "1 fois par jour"]
# valeurs / note par catégorie
note =np.array([0, 1, 2, 3, 4])
# nouvelles valeurs
note_new =np.array([0, 1, 2, 4, 3])
print("")
print("#####")
print("Question: 22...")
for ii in range(len(categ)):
    print("valeur: ",note[ii]," nouvelle valeur:",note_new[ii]," réponse: ",categ[ii])
# on conserve les infos dans la liste L
L.append([21,"22...",categ,note,note_new])

#####
# Début du preprocessing
#####

# Lecture des csv
# on utilise l'option sep=";" car ce sont des point virgules
# qui séparent les colonnes dans le fichier .csv
data0= pd.read_csv("data_eleonore.csv",sep=";")
data1= pd.read_csv("data_chiara.csv",sep=";")

# on vire les incomplets et ceux qui ne rentrent pas dans l'étude
data = data0.drop([0,3,14,27,29,37])
data_p = data1.drop([9,11])

# on garde seulement les colonnes du QUINOA
data_e = data.iloc[:, 4:26]
data_c = data_p.iloc[:, 7:29]

dne = data_e.to_numpy()
dnc = data_c.to_numpy()
alld = np.concatenate((dne,dnc))

# on concatène les tableaux pour n'en former qu'un
alldata = pd.concat([data_e,data_c],axis=0)
data = pd.DataFrame(alld)

#data = alldata.copy()

# on itère sur les colonnes du tableau
for it,col in enumerate(data):
    print("colonne numéro:",it)
    print("#####")
    print("Nom de la colonne: ",col)
    print("#####")

    val, categ = pd.factorize(data[col])

    # modification des valeurs en fonction de l'échelle définie au dessus

    # modification de data_new en données quantitatives
    nnew = L[it][4]
    data[col]=nnew[val]

```

```

# on affiche la réponse et la valeur qui correspond
for ii in range(len(categ)):
    print("valeur: ",ii,"\\t",valeur_new[ii],"\\t",réponse:\\t",categ[ii])

# on renomme les colonnes sous la forme Q#numéro
colna = data_e.columns
colna_new = []
for it, col in enumerate(colna):
    na = "Q{}".format(it+1)
    colna_new.append(na)
# changement de noms
data.columns = colna_new

# On rajoute une colonne au tableau, la colonne «Classe»
# qui est égale à 0 pour les donnée d'Éléonore et à 1 pour celle de Chiara
n1, m1 = data_e.shape
n2, m2 = data_c.shape
cl = np.concatenate((np.zeros(n1), np.ones(n2)))
# On créer la nouvelle colonne
data["Classe"] = cl

# Arrivé ici, normalement le tableau data ne contient plus que des données
# quantitatives (je n'ai pas vérifié)
data.describe()

# on sauvegarde le tableau obtenu en csv
data.to_csv("output.csv")

```

Annexe IX. Script d'apprentissage de la régression logistique et de la régularisation l1

```
# -*- coding: utf-8 -*-

import pandas as pd
import numpy as np
t=pd.read_csv("output.csv")
SEED = 30
from sklearn.model_selection import train_test_split
train_set, test_set = train_test_split(t, test_size=0.15, random_state=SEED)
#on sépare les entrées (x) et les sorties(y) :
ytrain=train_set["Classe"]
xtrain=train_set.drop("Classe", axis=1)
# On supprime la 1ère colonne qui numérotait les participations
xtrain=xtrain.drop("Unnamed: 0", axis=1)
#on fait la même chose pour la partie test :
ytest=test_set["Classe"]
xtest=test_set.drop("Classe", axis=1)
xtest=xtest.drop("Unnamed: 0", axis=1)

#on va faire l'apprentissage : régression logistique avec
# avec pénalité l1
# paramètre de pénalisation C : plus grand pour peu de pénalisation
# plus petit : plus de pénalisation
# Cp doit être un nombre >0
#Cp = 5.0
Cp = 0.7
#Cp = 4
print("Régression Logistique avec pénalisation l1, C = ",Cp)

from sklearn.linear_model import LogisticRegression
model = LogisticRegression(penalty='l1',C=Cp,solver='liblinear')
#ligne d'apprentissage :
model.fit(xtrain, ytrain)
#afficher le score de l'apprentissage:
print("Score", model.score(xtrain, ytrain))
#score = 100% donc il ne s'est jamais trompé
#on fait la même chose avec le test pour voir s'il arrive à généraliser son apprentissage :
print("Score test", model.score(xtest, ytest))
# score = 100% donc avec le test il ne se trompe pas. On va maintenant calculer la sensibilité et la spécificité :
from sklearn.metrics import precision_score, recall_score
model_pred=model.predict(xtrain)
se_train = recall_score(ytrain, model_pred)
sp_train = precision_score(ytrain, model_pred)
print("Sensibilité ensemble d'entraînement :", se_train)
print("Spécificité ensemble d'entraînement :", sp_train)
model_pred=model.predict(xtest)
se_test = recall_score(ytest, model_pred)
sp_test = precision_score(ytest, model_pred)
print("Sensibilité ensemble de test :", se_test)
print("Spécificité ensemble de test:", sp_test)

print("Poids du modèle : ")
print(model.coef_[-1])

print("biais b / intercept : ")
print(model.intercept_)

W=model.coef_[-1]
beta= model.intercept_[0]

for k in range(9):
    print("k :",k)
    scor= sum(xtest.iloc[k,:])
    print("Score :",scor)
    SRL = xtest.iloc[k,:]*W+beta
    print("Score RL:",SRL)
    print("Proba TAP pour RL:",np.exp(SRL)/(1+np.exp(SRL)))
    print("vraie y :",ytest.iloc[k])
```

Annexe X. Script ACP

```
# -*- coding: utf-8 -*-
"""
Created on Mon Apr 14 16:37:19 2025

@author: PC
"""

t<-read.csv("output2.csv")
t1<-t[1:57,2:25]
library("FactoMineR")
PCA(t1,scale.unit=F,quanti.sup=c(23,24))
```

Annexe XI. Script courbes ROC

```
# -*- coding: utf-8 -*-
"""
Created on Mon Apr 14 16:43:17 2025

@author: PC
"""

score_E_train = np.zeros(48)
for k in range(48):
    scor = sum(xtrain.iloc[k,:])
    score_E_train[k] = scor

# Roc curves
from sklearn.metrics import roc_curve

# We need the false positive rate and the true positive rate wrt thresholds
fpr, tpr, thresholds = roc_curve(ytrain ,model.decision_function(xtrain))
fpr2, tpr2, thresholds2 = roc_curve(ytrain , score_E_train)

# Function for plotting a nice ROC curve from the book of A. Géron
# [Hands-on machine learning with scikit-learn and Tensorflow]
def plot_roc_curve(fpr, tpr, label=None):
    plt.plot(fpr, tpr, linewidth=2, label= label)
    plt.plot([0,1], [0,1], 'k--')
    plt.axis([0,1,0,1])
    plt.xlabel('False Positive Rate')
    plt.ylabel('True Positive Rate')

# plot the curves
plot_roc_curve(fpr, tpr)
plot_roc_curve(fpr2, tpr2)
plt.title("ROC - Test set")
plt.show()
```

Annexe XII. Graphiques des résultats obtenus à chaque question du QUINOA

Question 1 : L'enfant bénéficie-t-il d'une rééducation (orthophonique, kinésithérapie, etc.) actuellement ?

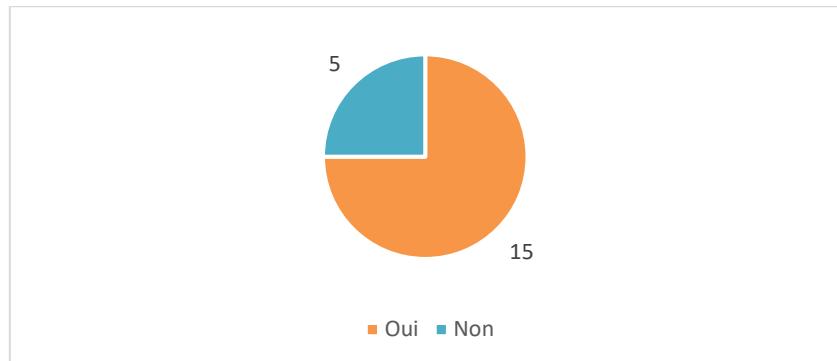


Figure 10 : Réponses à la question 1

Question 1bis : De quelle rééducation s'agit-il ?



Figure 11 : Réponses à la question 1bis

Question 1ter : Date de début de la rééducation

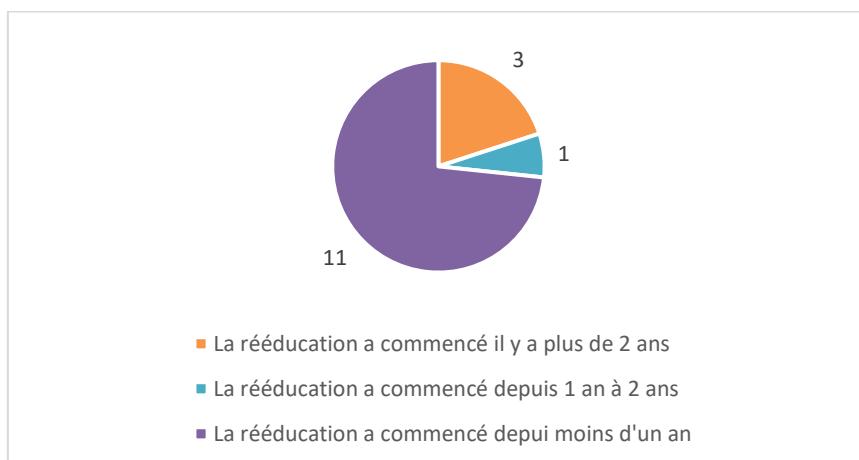


Figure 12 : Réponses à la question 1ter

Question 2 : Date de naissance

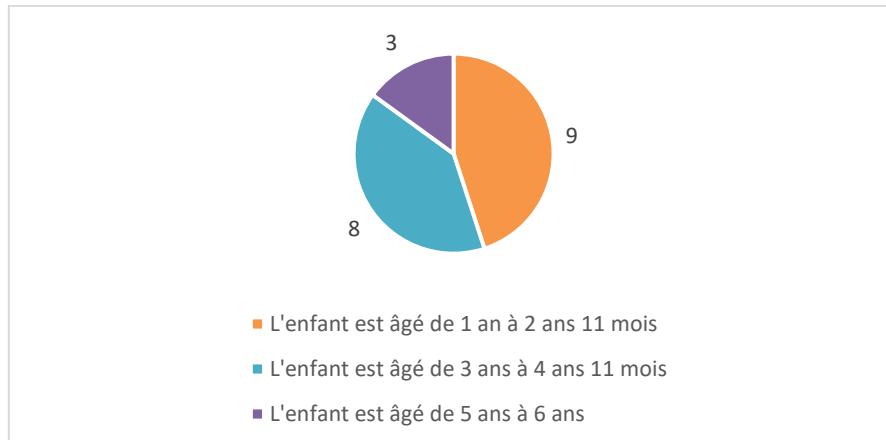


Figure 13 : Réponses à la question 2

Calcul de l'IMC grâce aux réponses aux questions 2 (date de naissance), 3 (poids) et 4 (taille).

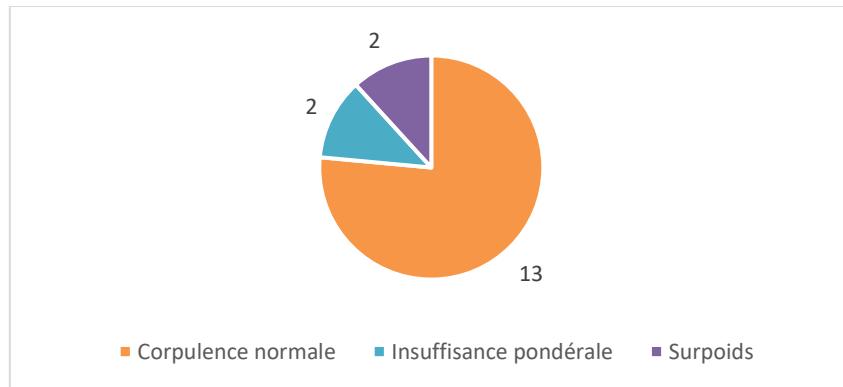


Figure 14 : Calcul de l'IMC selon les réponses aux questions 2, 3 et 4

Question 5 : « Lèvres/lange/joues/mâchoire de l'enfant sont-elles toniques ? »

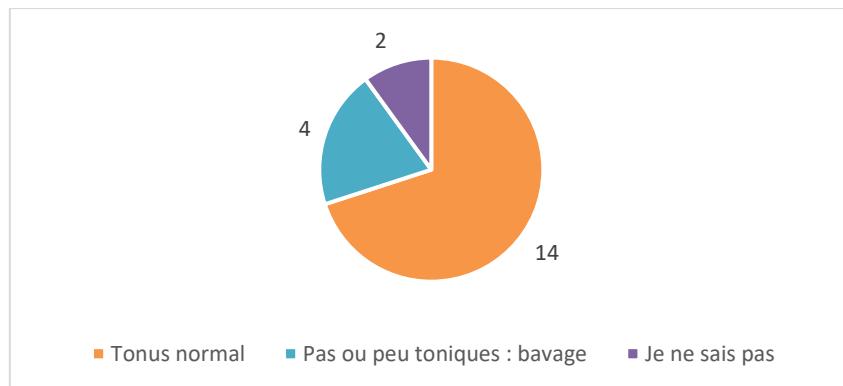


Figure 15 : Réponses à la question 5

Questions 6 : « L'enfant a-t-il (ou a-t-il eu) : »

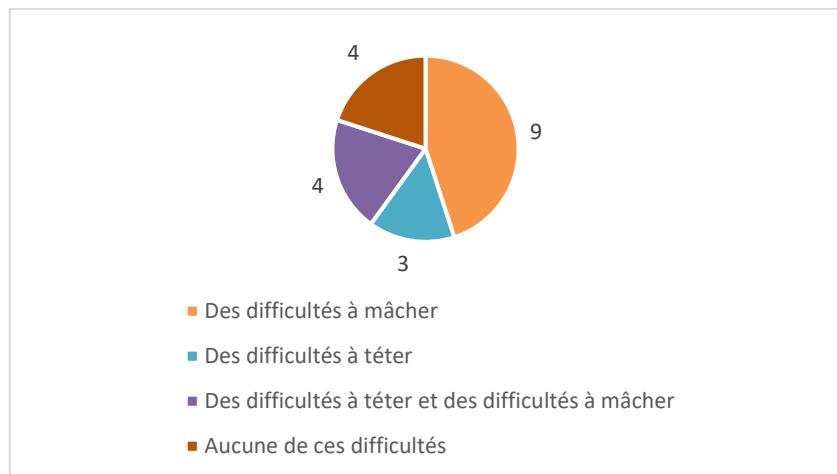


Figure 16 : Réponses à la question 6

Questions 7 : « L'enfant a-t-il (ou a-t-il eu) besoin d'interrompre la prise alimentaire parce qu'il a (ou avait) l'air fatigué ou en difficulté ? »

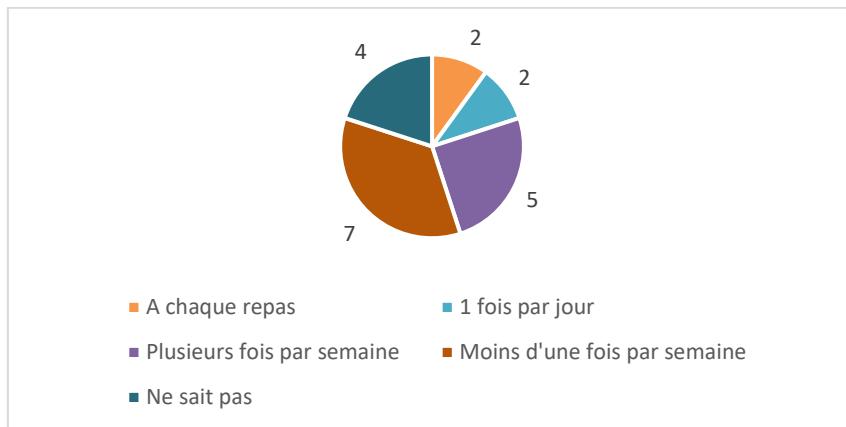


Figure 17 : Réponses à la question 7

Question 8 : L'enfant accumule-t-il de la nourriture dans ses jours et/ou la recrache après l'avoir stockée ?

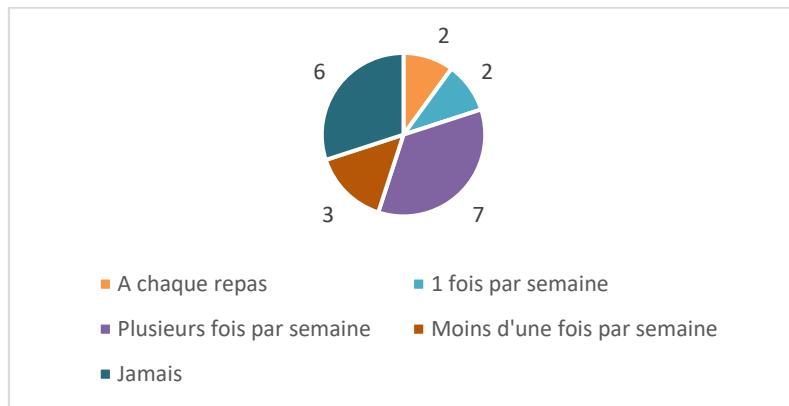


Figure 18 : Réponses à la question 8

Question 9 : « En général, le repas de l'enfant dure en moyenne : »

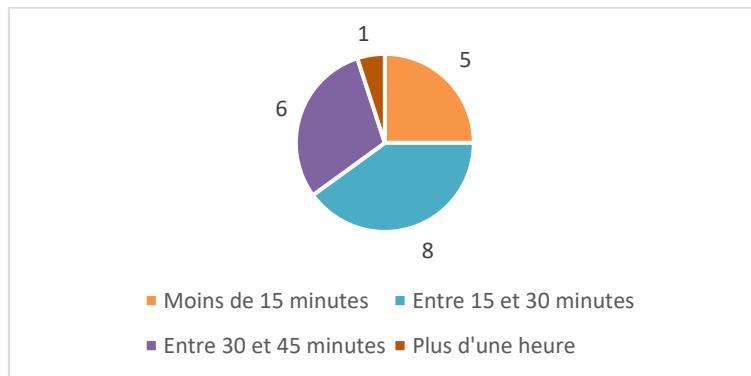


Figure 19 : Réponses à la question 9

Question 10 : « Estimez-vous que votre enfant est plus « casse-cou » que les autres enfants de son âge ? »

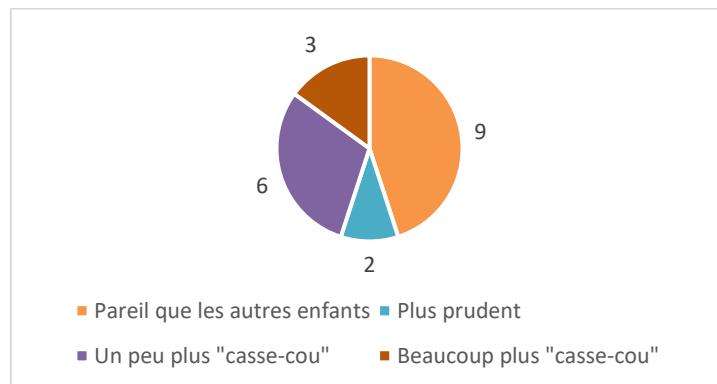


Figure 20 : Réponses à la question 10

Question 11 : « A quelle fréquence l'enfant : - refuse qu'on lui touche le visage ou le corps (caresses, bain, câlins) ou – refuse le contact des mains, des pieds ou du corps avec certaines matières (aliments, textiles, herbe, sable, pâte à modeler...) »

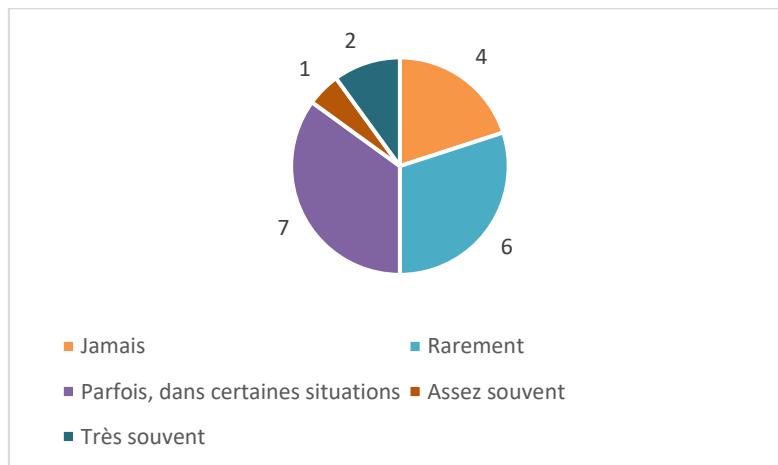


Figure 21 : Réponses à la question 11

Question 12 : « L'enfant porte-t-il (ou a-t-il porté) les objets ou ses doigts à la bouche en dehors des périodes de poussées dentaires ? »

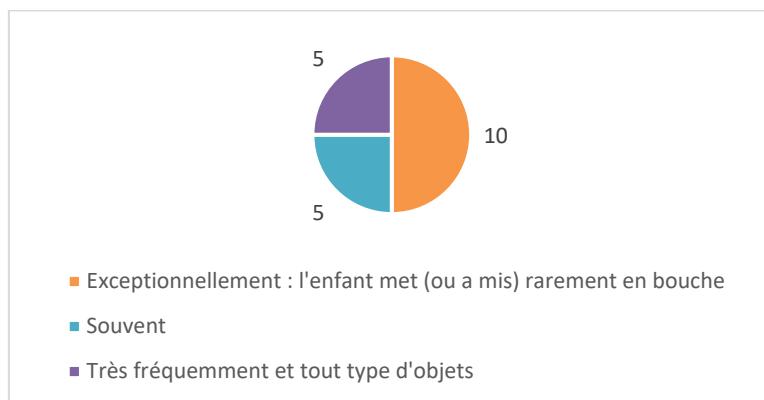


Figure 22 : Réponses à la question 12

Question 13 : « Lors du repas ou de l'approche d'un objet ou aliment (en dehors des aliments favoris ou acceptés depuis toujours) de la bouche, à quelle fréquence l'enfant manifeste-t-il l'un ou plusieurs de ces comportements ? »

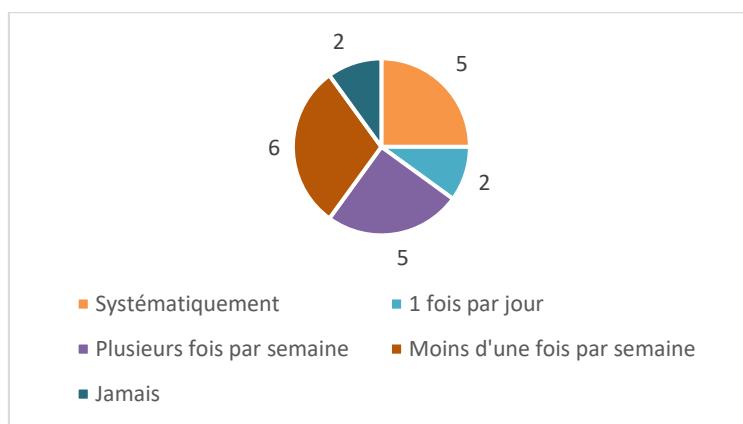


Figure 23 : Réponses à la question 13

Question 14 : « A quelle fréquence l'enfant a-t-il des haut-le-cœur lors de l'introduction d'un aliment (en dehors des aliments favoris ou acceptés depuis toujours) ou d'un objet en particulier dans sa bouche ? »

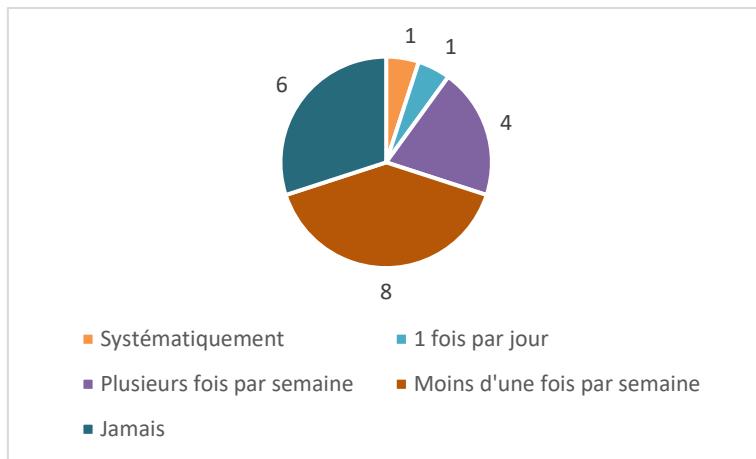


Figure 24 : Réponses à la question 14

Question 15 : « Diriez-vous que l'introduction de nouvelles textures (passage aux purées ou passage aux morceaux » a été :

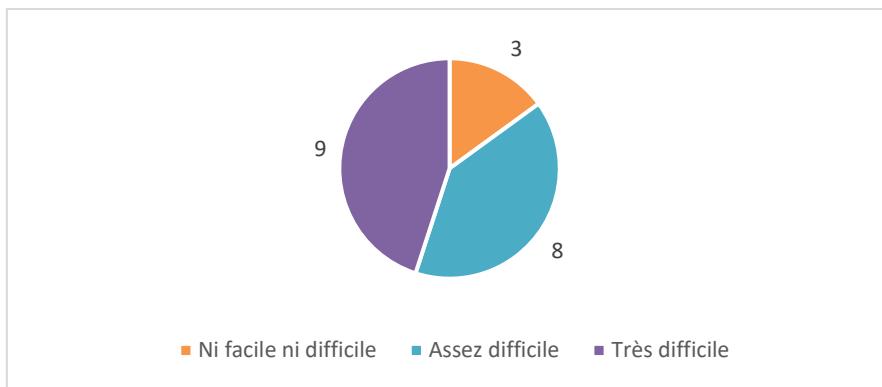


Figure 25 : Réponses à la question 15

Questions 16 : « L'enfant refuse-t-il de goûter des aliments nouveaux, choisit les aliments qu'il peut manger et trie son assiette ? »

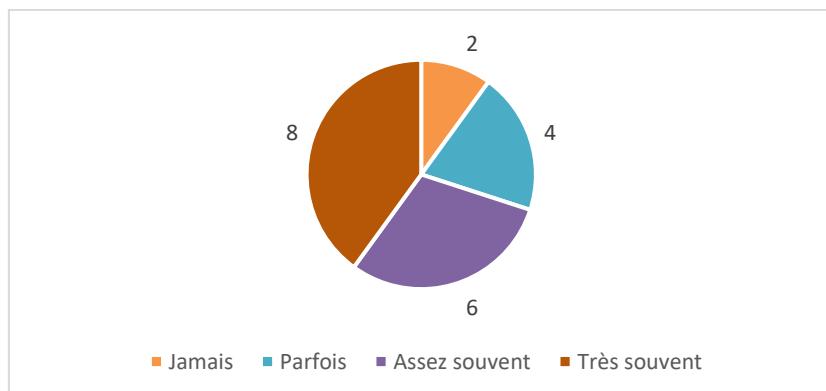


Figure 26 : Réponses à la question 16

Question 17 : « L'enfant a-t-il une préférence pour les aliments au goût prononcé ou l'enfant prend-il une grande quantité d'aliments en bouche ? »

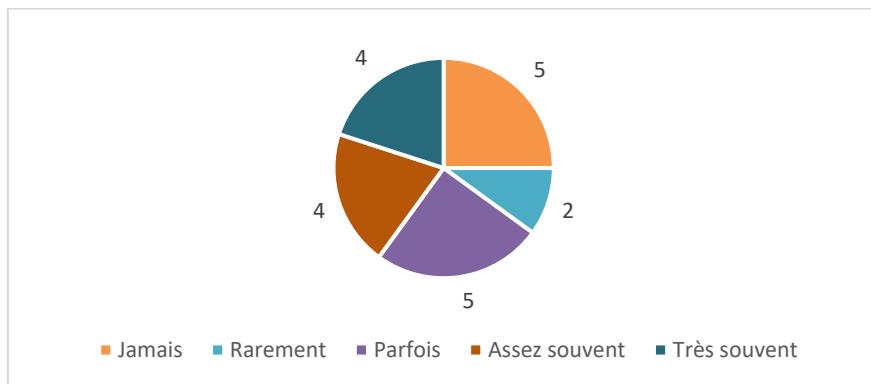


Figure 27 : Réponses à la question 17

Question 18 : « L'enfant manifeste-t-il sa faim ? »

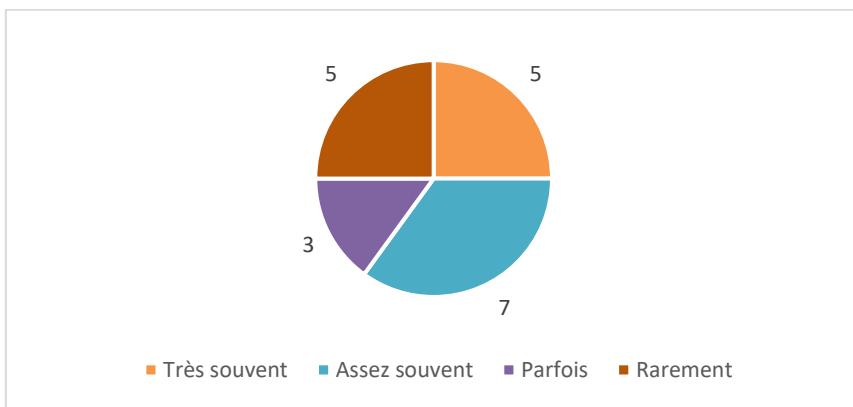


Figure 28 : Réponses à la question 18

Question 19 : « L'enfant redoute-t-il les moments de repas, le passage à table se fait-il difficilement ? »

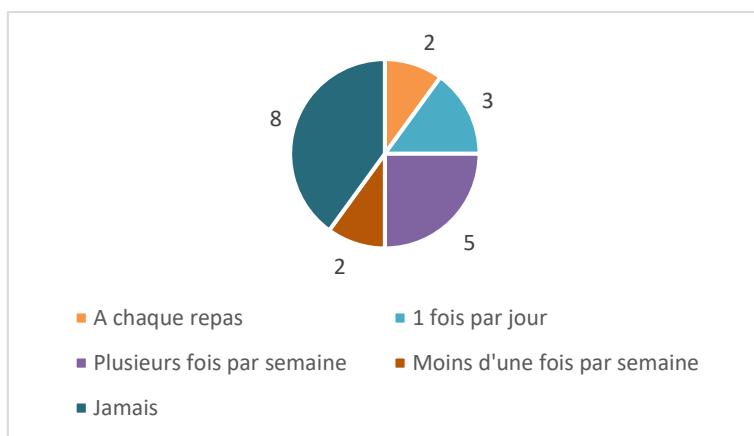


Figure 29 : Réponses à la question 19

Question 20 : « L'enfant a-t-il des attitudes de défense, d'évitement envers les aliments (en dehors des aliments favoris ou acceptés depuis toujours) ? »

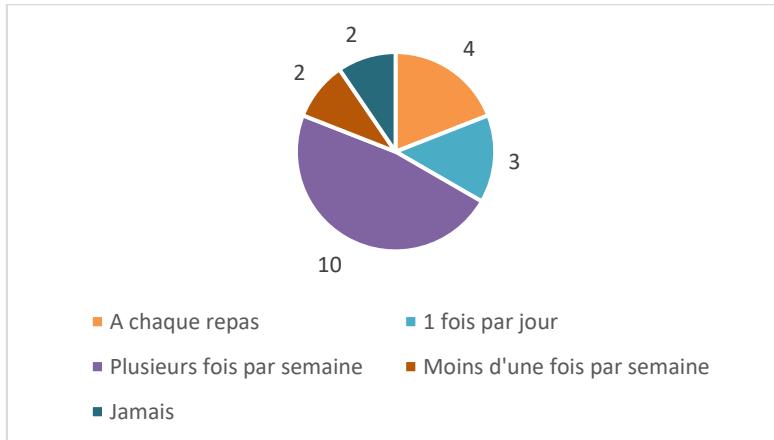


Figure 30 : Réponses à la question 20

Question 21 : « L'enfant réagit-il différemment (refus de manger...) selon la personne avec qui il mange ? »

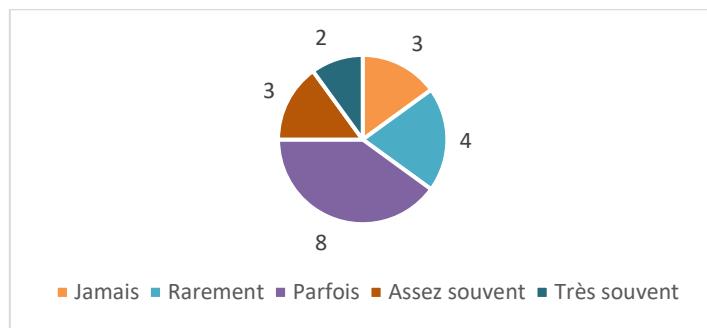


Figure 31 : Réponses à la question 21

Question 22 : « L'enfant a-t-il peur d'avaler et/ou de s'étouffer ? »

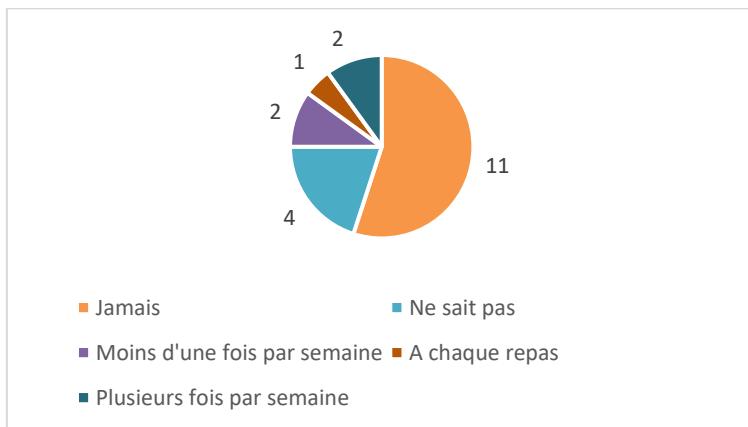


Figure 32 : Réponses à la question 22

Question 23 : « Les repas sont-ils difficiles, sources de conflits et d'anxiété pour vous et votre enfant ? »

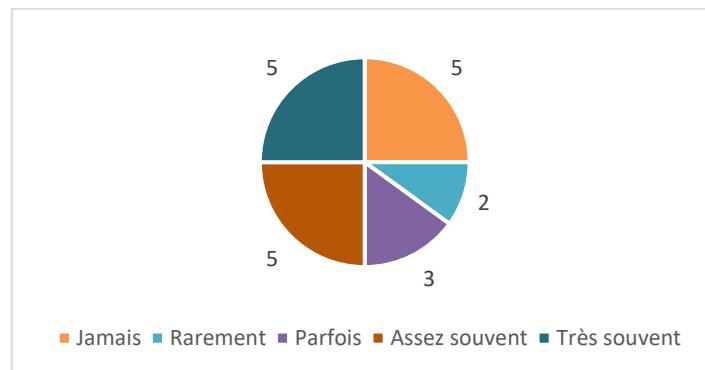


Figure 33 : Réponses à la question 23

Question 24 : « Utilisez-vous des stratégies dans le but que votre enfant mange ? »

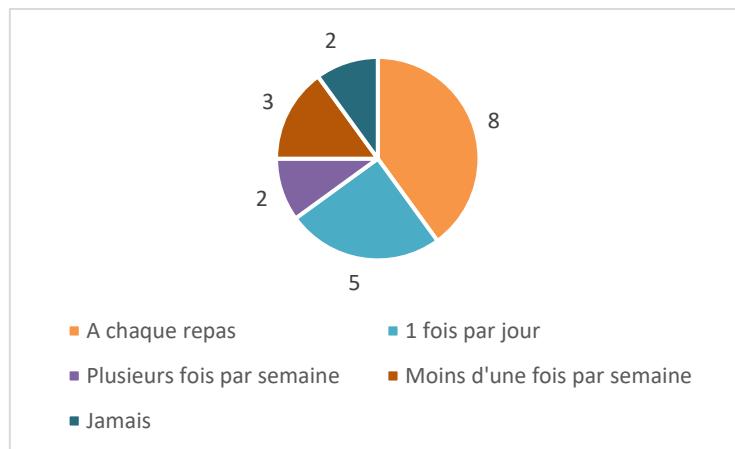


Figure 34 : Réponses à la question 24

Question 25 : « L'enfant fait-il des fausses-routes (la nourriture passe de travers) ? »

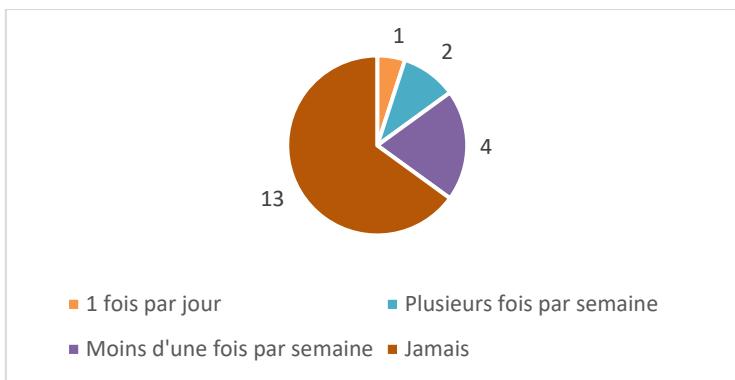


Figure 35 : Réponses à la question 25

Question 26 : « Enrichissez-vous l'alimentation de votre enfant parce que vous avez l'impression qu'il ne mange pas assez ? »

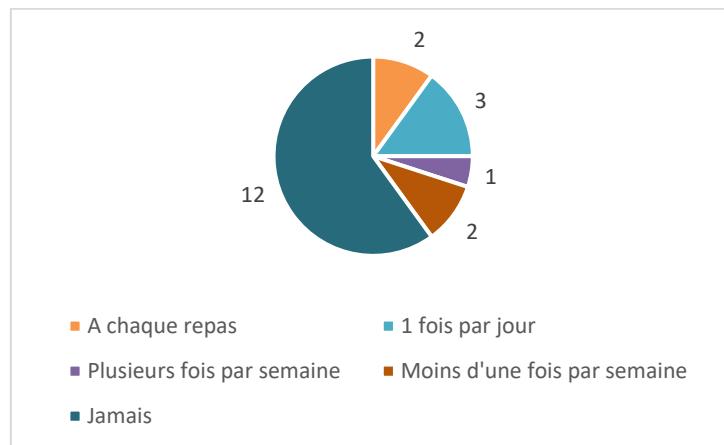


Figure 36 : Réponses à la question 26

Annexe XIII. Poids des questions de la régression logistique

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Poids	3,22E-04	9,02E-01	4,27E-01	4,00E-01	2,88E-01	-1,10E-01	4,24E-01	2,00E-01	4,83E-01	2,75E-01

Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
8,57E-01	4,79E-01	4,23E-01	-4,26E-02	1,14E-01	4,05E-01	1,86E-01	1,48E-01	2,71E-01

Q20	Q21	Q22
4,86E-01	3,83E-01	4,20E-01

Annexe XIV. Poids des questions de la régression logistique avec la régularisation l1

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Poids	0	0.58912096	0.11653824	0	0	-0.07424676	0	0	0.08384847	0

Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19
0.88865588	0	0	0	0	0.35153658	0	0	0.22364791

Q20	Q21	Q22
0.24345331	0	0.4072142

L'apport des modèles d'apprentissage dans les étalonnages : Poursuite de l'étalonnage du QUINOA (QUestionnaire d'INvestigation de l'Oralité Alimentaire)

L'étalonnage est une étape clé dans la création et la validité d'un outil de décision. Il consiste en la passation d'un même test à deux populations : une saine et une porteuse d'un trouble. L'étalonnage permet d'établir une norme pour l'outil de décision en question. Nous avons détaillé le concept d'étalonnage à travers un exemple concret : le QUestionnaire d'INvestigation de l'Oralité Alimentaire (QUINOA). Il s'agit d'un test de dépistage des troubles alimentaires pédiatriques (TAP) chez un enfant tout-venant, âgé de 1 an à 6 ans, en création depuis 2018. La première partie de l'étalonnage sur une population saine a été effectuée en 2024, et notre travail a consisté à réaliser l'étalonnage sur une population porteuse d'un TAP. Nous avons reçu 20 questionnaires valides, qui nous ont permis d'aboutir à un étalonnage sensible et spécifique. Pour réaliser l'étalonnage, nous avons utilisé le modèle de régression logistique, qui est un modèle d'apprentissage supervisé permettant de réduire le risque d'erreurs. Face à l'échantillonnage restreint, il semblerait préférable de poursuivre l'étalonnage, afin qu'il soit plus représentatif de la population. La mise à disposition du QUINOA auprès des professionnels de santé permettrait que l'enfant soit dirigé au plus tôt chez un orthophoniste afin de débuter une prise en soin et ainsi éviter la chronicisation des troubles et limiter leurs conséquences.

Mots-clés : étalonnage, modèles d'apprentissage, statistiques, troubles alimentaires pédiatriques, dépistage

The contribution of learning models to calibration: Continued calibration of QUINOA

Calibration is a key step in the creation and validity of a decision tool. It consists of administering the same test to two populations: one healthy group and one with a disorder. Calibration establishes a standard for the decision tool in question. We have detailed the concept of calibration through a concrete example: the QUINOA. This is a screening test for pediatric feeding disorders in a general population of children aged between 1 and 6 years, in creation since 2018. The first part of the calibration on a healthy population was carried out in 2024, and our work consisted in making the calibration on a population with pediatric feeding disorders. We received 20 valid questionnaires, which enabled us to achieve a sensitive and specific calibration. To perform calibration, we used the logistic regression model, which is a supervised learning model that reduces the risk of errors. Faced with limited sampling, it would seem preferable to continue the calibration, so that it is more representative of the population. Making the QUINOA available to healthcare professionals would allow the child to be referred as early as possible to a speech therapist to begin care, thereby avoiding the chronicity of disorders and limiting their consequences

Keywords : calibration, learning model, statistics, pediatric feeding disorders, screening

