Validité et fiabilité du MABC-2 et du DCDQ dans le diagnostic des Troubles Développementaux de la Coordination

Hugo Bourniquel

▶ **To cite this version:**

Hugo Bourniquel. Validité et fiabilité du MABC-2 et du DCDQ dans le diagnostic des Troubles Développementaux de la Coordination. Médecine humaine et pathologie. 2020. dumas-03079264

HAL Id: dumas-03079264

https://dumas.ccsd.cnrs.fr/dumas-03079264

Submitted on 17 Dec 2020

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

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

AIX-MARSEILLE UNIVERSITÉ
ÉCOLE DES SCIENCES DE LA RÉADAPTATION
FORMATION EN MASSO-KINÉSITHÉRAPIE

Validité et fiabilité du MABC-2 et du DCDQ dans le diagnostic des Troubles Développementaux de la Coordination

BOURNIQUEL Hugo

Directeur de mémoire : Mme Béatrice CAORS

D.E.M.K MARSEILLE 2020

Remerciements :

Je tiens à remercier dans un premier temps Madame Caors, sans qui ce mémoire n'aurait pas été possible. A son attention à mon égard, à sa passion de la kinésithérapie et de l'enseignement, à ses conseils avisés professionnels comme personnels pendant 4 ans ainsi que son intérêt stimulant porté à mon projet.

Je voudrais ensuite adresser un remerciement pour Jean François Connan, mon psychomotricien depuis 3 ans maintenant, qui m'a amené à réaliser ce travail en m'inspirant sur le traitement des troubles moteurs du développement, qui m'a offert la possibilité en tant qu'adulte de redécouvrir une nouvelle manière de réfléchir, une façon d'apprivoiser son corps et le mouvement ce qui m'a permis de m'épanouir dans ma formation.

Un immense merci à ma famille, mes parents, mon frère, mes grands-parents, pour leur soutien sans faille tout au long de ma vie, notamment pendant mes six longues années d'études supérieures, malgré les efforts et les sacrifices que cela a parfois pu engendrer, la patience que je nécessite souvent et pour tout l'amour et la confiance qu'ils ont su m'apporter toujours.

Enfin, un merci particulier pour mes amis qui ont accepté ma singularité et plus spécialement mon binôme, mon alter égo Théo qui me fait croire au destin et qui a rendu unique et inoubliable mon expérience dans cette école mais aussi tous les moments partagés ici et ailleurs. A nos futures péripéties…

Lexique :

TDC : Troubles Développementaux de la Coordination

APA : Association Psychiatrique Américaine

DCDQ : Questionnaire sur le Trouble Développemental de la Coordination

MABC-2 : Movement Assessment Battery for Children - second edition, deuxième édition de la batterie d'évaluation motrice pour les enfants

Se : Sensibilité

Sp : Spécificité

ASC : Aire sous la Courbe

VPP : Validité prédictive positive

VPN : Validité prédictive négative

DSM-V : 5ème version du Diagnostic and Statistical Manuel of Mental Disorders

MK : Masseur Kinésithérapeute

BOT-2 : Bruininks-Oseretsky Test of Motor Proficiency, deuxième édition

CSAPPA : the Children's Self-perceptions of Adequacy in and Predilection for Physical Activity

BOTMP : Bruininks-Oseretsky Test of Motor Proficiency

r : coefficient de corrélation de Pearson

α : alpha de Cronbach

ρ : coefficient de Spearman

INSERM :

HAS : Haute Autorité de Santé

ICC : Corrélation intra classe

IC : Intervalle de Confiance

1 Introduction 1
1.1 Les Troubles Développementaux de la Coordination 1
1.1.1 Définition 1
1.1.2 Physiopathologie 2
1.1.2.1 Physiologie 2
1.1.2.2 Critères diagnostiques du TDC (DSM-V, 2013) 3
1.1.2.3 Symptômes 3
1.1.2.4 Causes et facteurs de risque 4
1.1.3 Les outils diagnostics 5
1.1.3.1 Le MABC-2 6
1.1.3.2 Le DCDQ 8
1.1.4 Traitement 9
1.1.4.1 Méthodes 9
1.1.4.2 Imagerie motrice 9
1.1.4.3 NTT (Neuromoteur Task Training) 10
1.1.4.4 CO-OP (Cognitive Orientation to daily Occupational Performance) 10
1.1.5 Intérêts de cette revue de littérature 11
1.1.6 Objectifs de cette revue de littérature 13
2 Méthodologie 14
2.1 Critères d'éligibilité des études pour cette revue 14
2.1.1 Types d'études 14
2.1.2 Population/Pathologie 15
2.1.3 Outil de mesure 16
2.1.4 Objectif/critères de jugement 16
2.2 Méthodologie de recherche des études 16
2.2.1 Sources documentaires investiguées 16
2.2.2 Equation de recherche utilisée (mots clefs) 17
2.3 Extraction et analyse des données 17
2.3.1 Sélection des articles 17
2.3.2 Extraction des données 18
2.3.3 Evaluation de la qualité méthodologique des études sélectionnées 18
2.3.4 Méthode de synthèse des résultats 19
3 Résultats 20

3.1 Description des études 20
3.1.1 Résultats de la recherche bibliographique - sélection des études 20
3.1.2 Diagramme de flux 21
3.1.2.2 Sélection 21
3.1.2.3 Inclusion 21
3.1.2.3 Eligibilité 21
3.1.2.1 Identification 21
3.1.3 Caractéristiques des études exclues 21
3.1.4 Caractéristiques des études incluses 22
3.1.5 Tableau synthétique 24
3.2 Risques de biais des études incluses 28
3.2.1 Grille d'analyse utilisée 28
3.2.2 Synthèse des biais retrouvés 28
3.3 Effets de l'intervention 31
3.3.1 Ferreira L et al, 2019 31
3.3.2 Martini R et al, 2011 31
3.3.3 Montoro A et al, 2016 32
3.3.4 Capistrano R et al, 2015 33
3.3.5 Schoemaker M et al, 2012 34
3.3.6 Asunta P et al, 2019 35
3.3.7 Ray-Kaeser S et al, 2019 35
3.3.8 Wilson B et al, 2009 36
3.3.9 Jaikaew R et al, 2019 37
3.3.10 Ellinoudis T et al, 2011 37
3.3.11 Wuang Y et al, 2012 38
3.3.12 Pannekoek L et al, 2012 38
3.3.13 Tseng M et al, 2012 39
3.3.14 Rivard L et al, 2014 39
3.3.15 Cairney J et al, 2008 39
3.3.16 Caravale B et al, 2015 40
3.3.17 Caravale B et al, 2014 40
3.3.18 Holm I et al, 2013 41
4 Discussion 42

4.1 Analyse des principaux résultats 42
4.2 Applicabilité des résultats en pratique clinique 50
4.3 Qualité des preuves 51
4.4 Biais potentiels de la revue 53
5 Conclusion 55
6 Bibliographie 57

1 Introduction

Selon l'INSERM [1], malgré une difficulté à estimer réellement cette pathologie, 5 à 7% de la population française en souffrirait à des degrés variables avec une prédominance masculine marquée [2]. Vivre avec une dyspraxie, en 2019, en France, peut représenter un parcours du combattant et les réponses apportées aux enfants dyspraxiques et à leurs familles sont loin de donner satisfaction. Le flou statistique constaté peut s'expliquer par plusieurs facteurs (difficultés de diagnostic, appellation différente selon les médecins, degré d'atteinte qui ne permettent pas forcément de parler de dyspraxie en tant que telle), acceptation sociale non optimale (mon enfant n'a pas de problème, vécu comme une excuse par l'encadrement). La dyspraxie est très souvent pensée comme une pathologie pour une population d'enfant mais elle peut toucher des adultes non diagnostiqués ou non traités efficacement. C'est un trouble méconnu et il peut être considéré comme un handicap invisible ce qui rend délicat son diagnostic.

1.1 Les Troubles Développementaux de la Coordination

1.1.1 Définition

Le mot dyspraxie vient du grec dys « difficulté, mauvais fonctionnement » et du grec praxis « action ». Une praxie est la coordination de l'activité gestuelle. Depuis juin 2015, la dénomination internationale est dorénavant le trouble du développement de la coordination (TDC), la communauté internationale ayant accepté la nouvelle terminologie proposée par le DSM-V, dernière version du Diagnostic and Statistical Manuel of Mental Disorders de 2013, livre de référence des troubles mentaux de l'Association américaine de psychiatrie. Ce terme générique permet une plus grande uniformité mondiale de l'appellation, une meilleure compréhension et une meilleure intégration de la diversité des troubles que cela engendre. Le Trouble développemental de la coordination (TDC), selon les dernières définitions, se définit par un retard dans les apprentissages moteurs et dans la coordination des mouvements en comparaison aux autres enfants du même âge. Les difficultés motrices qui en résultent rendent la réalisation des tâches quotidiennes plus ardue pour l'enfant qui présente un TDC. Dans la Classification Internationale des Maladies 10e révision (CIM10), les termes TDC et dyspraxie de développement sont considérés comme équivalents. Mais il reste encore compliqué de séparer les termes internationaux de TDC qui recouvrent une grande hétérogénéité de troubles, de la dyspraxie ou encore des Troubles d'Acquisition de la Coordination (TAC). Il existe encore un défaut d'harmonisation. Le CIM10, gérée par l'Organisation Mondiale de la Santé, classifiait ainsi la dyspraxie en trouble cognitif spécifique des apprentissages et plus exactement en trouble spécifique du développement moteur. Le DSM-V classifie quant à lui les TDC comme un trouble neurodéveloppemental dans la catégorie moteur [3]. C'est ainsi une pathologie sans antécédent neurologique à l'inverse de l'enfant paralysé cérébral par exemple [4].

1.1.2 Physiopathologie

1.1.2.1 Physiologie

Le TDC est un trouble neurodéveloppemental et un trouble qui repose sur des diagnostics indiquant la présence chez l'enfant d'un trouble développemental se manifestant principalement dans des habiletés gestuelles, et se traduisant souvent par des difficultés d'apprentissage à l'école ainsi que par des limitations dans d'autres activités de la vie quotidienne que ce soit à la maison ou par exemple dans les activités extra scolaires sportives ou manuelles (la personne est jugée lente, maladroite, peu habile). Cette maladie se caractérise par une affection de la planification des mouvements et la coordination en raison d'une altération de la communication entre le cerveau et le corps et se distingue en particulier par une altération de la capacité à exécuter de manière automatique des mouvements déterminés, en l'absence de toute paralysie ou parésie des muscles impliqués dans le mouvement. Le sujet doit contrôler volontairement chacun de ses gestes, ce qui est très coûteux en attention, et rend la coordination des mouvements complexes de la vie courante extrêmement difficile, donc rarement obtenue [5].

Un geste est un ensemble de mouvements coordonnés dans le temps et l'espace en vue de la réalisation d'une action volontaire, finalisée. Les TDC peuvent se définir comme étant des anomalies de constitution des gestes et de l'organisation visuo-spatiale ou comme étant une pathologie de la conception et/ou de la planification–programmation donc de la réalisation des gestes appris.

A la naissance, le cerveau de l'enfant est doté d'un ensemble de solutions qui lui permettra de coordonner et planifier, programmer des gestes volontaires après une période d'apprentissage. Ces gestes finiront par être automatisés. L'automatisation permet la double tâche notamment en s'appuyant sur un certain nombre de repères visuo-spatiaux (repères dans l'espace, situer les éléments les uns par rapport aux autres). La dyspraxie est par conséquent un trouble cognitif spécifique mais comme les autres troubles DYS, la dyspraxie ne remet absolument pas en question l'intelligence de la personne, il se peut même qu'elle soit associée à une précocité intellectuelle.

L'enfant dyspraxique va progresser dans l'apprentissage de certains gestes mais d'autres resteront difficiles voire impossibles à réaliser. Chaque enfant dyspraxique étant différent, les apprentissages ne se feront pas à la même vitesse, et si certains enfants parviendront à automatiser certains gestes, d'autres n'y parviendront jamais. A des degrés plus ou moins sévères, le TDC est un trouble durable et persistant avec des répercussions dans la vie quotidienne, à l'école, dans l'accès à la vie sociale, les loisirs, et l'emploi. En fonction de ces répercussions, le TDC crée alors une situation de handicap nécessitant des adaptations et/ou compensations [6].

Elle entraîne alors fatigue et une certaine lenteur dont il faut tenir compte au quotidien et dans les différents apprentissages tout au long de la vie mais cela entraîne souvent également une mésestime de soi car le résultat ne sera pas forcément à la hauteur de l'effort consenti (discordance entre l'acte voulu et l'acte réalisé) et l'enfant ou l'adulte concerné, en sera pleinement conscient. Donc on s'appuiera sur ses points forts pour l'aider dans sa progression.

Il est connu qu'il existe plusieurs types de dyspraxie :

- Les dyspraxies constructives avec une difficulté à assembler des pièces pour construire un tout. Le problème est dans l'assemblage des pièces les unes par rapport aux autres (bricolage, couture, cubes, puzzles, …)
- La dyspraxie constructive visuo-spatiale qui se caractérise par une difficulté à se repérer dans l'espace (orientation, fixer une ligne pour lire, suivre la trajectoire d'un objet, rechercher une information dans un texte, lire un plan, géométrie, …)
- La dyspraxie idéatoire, c'est-à-dire une difficulté de réaliser un geste avec un objet ou un outil (gestes du quotidien avec par exemple hygiène, dessin, mécanique, ménage mais aussi des gestes simples comme allumer une allumette ou utiliser un tournevis, …)

Il existe également d'autres sous catégories de dyspraxie : non constructive, idéomotrice, d'habillage, dysgraphique et oro-faciale [7][8].

1.1.2.2 Critères diagnostiques du TDC (DSM-V, 2013)

Le manuel diagnostique et statistique des troubles mentaux (également désigné par le sigle DSM, abréviation de l'anglais : Diagnostic and Statistical Manual of Mental Disorders) est un ouvrage de référence publié par l'Association américaine de psychiatrie décrivant et classifiant les troubles mentaux. Le manuel évolue à partir des statistiques collectées dans la pratique clinique et les recherches scientifiques. Il a été radicalement révisé entre 1980 et la dernière édition, la cinquième qui est publiée en 2013. Il permet une certaine uniformité dans le diagnostic à l'échelle mondiale en donnant un cadre théorique aux pathologies et à leurs outils diagnostics. [9]

Critère A : l'acquisition et l'exécution d'habiletés motrices coordonnées sont nettement au-dessous du niveau escompté compte tenu de l'âge chronologique du sujet et en dépit d'occasions d'apprentissage et d'utilisation de ces habiletés. Les difficultés se traduisent par de la maladresse (laisser tomber ou heurter des objets) ainsi que de la lenteur et de l'imprécision dans l'exécution des habiletés motrices (attraper un objet, utiliser des ciseaux ou des couverts, écrire, faire du vélo, pratiquer une activité sportive)

Critère B : le déficit en habiletés motrices du critère A interfère de façon significative et persistante avec les activités de la vie courante appropriées à l'âge chronologique (soins et entretien de soi) et a des conséquences sur la réussite scolaire, les activités préprofessionnelles et professionnelles, les loisirs et les jeux

Critère C : le début des symptômes se situe dans la première enfance

Critère D : les difficultés motrices ne sont pas mieux expliquées par une déficience intellectuelle ou une déficience visuelle et ne sont pas attribuables à une condition neurologique qui affecte le mouvement.

1.1.2.3 Symptômes

Des symptômes peuvent renseignés sur un potentiel trouble et permettent d'envisager un TDC [1][7] :

- Troubles du développement moteur avec maladresse et difficulté à exécuter des mouvements volontaires et coordonnés (marche, bicyclette, nage, jeux d'adresse, manier ses couverts, s'habiller seul, se brosser les dents, nouer ses lacets, se laver seul, contrôler sa force) mais aussi difficultés à apprendre de nouveaux gestes ou aussi à la répétition des gestes
- Fatigabilité, concentration excessive pour une action apparaissant simple aux autres enfants
- Difficultés dans la motricité fine
- Stratégie d'évitement (billes, jeux de constructions, instrument de musique)
- Problèmes d'orientation dans l'espace ou dans l'analyse d'un texte par exemple
- Sous-utilisation possible des membres supérieurs du bébé
- Dysgraphie (écriture ou dessin), troubles oculaires, de la parole, orthophoniques, logico-mathématiques, de la mémoire
- Conséquences possibles : un manque de soin, une apparence physique négligée et un manque de confiance en soi dû au ressenti de l'enfant dyspraxique de sa différence. Les enfants dyspraxiques peuvent avoir une tendance à se refermer sur eux-mêmes ce qui peut avoir des conséquences sur la vie sociale. La scolarisation peut être difficile pour l'enfant dyspraxique qui peut se sentir rejeté par les remarques ou le rejet de ses professeurs et de ses camarades. Les dyspraxiques peuvent également avoir un raisonnement n'allant pas dans le même sens que celui des autres et développer, à la suite de ces difficultés motrices, une forme d'ambidextrie.
- L'organisation est un véritable souci pour les dyspraxiques. Parmi les conséquences pour les adultes de la maladie, des difficultés notoires concernant le rangement. Ces problèmes organisationnels peuvent être aperçus dans l'ensemble des tâches manuelles.

On peut constater des troubles récurrents chez un patient TDC :

- Troubles de la coordination et de l'exécution évidemment
- Tonus postural altéré (hypertonie, contractions involontaires)
- Troubles de l'équilibre
- Troubles de la dissociation
- Troubles de l'habilité
- Troubles de proprioception ou de perception sensorielle
- Difficultés pour contrôler sa force
- Difficultés du travail en endurance (fatigue et perte de qualité du geste)
- Défaut de visualisation et de temporalité
- Défaut d'automatisation, de l'intégration et l'analyse d'information et de planification
- Troubles de l'orientation et de gestion de l'espace (lui-même et l'environnement alentour)
- Défaut de gestion émotionnelle ou du stress (anxiété notamment)

1.1.2.4 Causes et facteurs de risque

Les causes des TDC ne sont pas clairement identifiées et seraient diverses. Les causes demeurent ainsi inconnues. Un facteur de risque est tout attribut, caractéristique ou

exposition d'un sujet qui augmente la probabilité de développer une pathologie. Certains facteurs seraient ainsi favorisants de TDC, comme la génétique qui est un facteur suspecté car dans de nombreuses familles, on retrouve plusieurs cas de " dys ". Mais aussi d'autres hypothèses comme un faible poids à la naissance, la prématurité, un accouchement dystocique avec anoxie ou une grossesse avec alcoolisme prénatal. Toutefois, selon l'institut des troubles d'apprentissage, le monde scientifique n'a pas trouvé de consensus actuellement sur l'origine des anomalies cérébrales liées au TDC. Les recherches se poursuivent en ce sens pour nous renseigner sur le sujet. Des études sont menées sur des possibles atteintes neurologiques comme l'absence de dominance cérébral, des difficultés de connexion entre les différentes régions cérébrales impliquées dans le geste, ou entre les deux hémisphères, une anomalie cérébrale diffuse : atrophie corticale, dilatation des ventricules, démyélinisation (absence de myéline, substance qui favorise la connexion entre les neurones), des troubles oculomoteurs ou neuro-visuels, c'est à dire de l'orientation du regard. Il pourrait également s'agir d'une atteinte du cervelet, l'organe nerveux responsable d'une grande partie de la coordination motrice. Si l'existence d'un fonctionnement cérébral différent chez les enfants présentant un TDC semble avérée, il est impossible au vu de ces travaux de déterminer précisément les régions qui sous-tendent ce trouble. En effet, les régions observées comme fonctionnant différemment sont multiples. Le cervelet et la région pariétale ont été identifiés dans plusieurs études mais on retrouve aussi l'implication des régions frontales, des ganglions de la base, du cortex cingulaire ou du gyrus lingual. Par ailleurs, le faible nombre d'études et leur qualité ne permettent pas de conclure de façon définitive, ni d'infirmer ou de valider les hypothèses sur les corrélats neuronaux qui coexistent dans la littérature, notamment concernant l'implication du cervelet, du cortex pariétal postérieur et du cortex frontal inférieur ou encore de la voie dorsale (occipito-pariétale) ou des boucles corticocérébelleuse et cortico-striatale. Les résultats ne permettent pas non plus d'identifier l'origine du trouble et ses conséquences sur les autres structures cérébrales. Le TDC est ainsi parfois accompagné par d'autres troubles développementaux : un trouble déficitaire de l'attention avec ou sans hyperactivité (TDAH), des troubles d'apprentissages, un trouble du langage, un trouble du spectre de l'autisme, une déficience intellectuelle ou encore de l'épilepsie. La neuro-imagerie n'est pas un outil de diagnostic du TDC [1]. Ainsi, devant l'absence de preuve physiologique, le rôle du médecin est d'exclure d'autres diagnostics neurologiques, métaboliques, sensoriels ou génétiques. Les recommandations européennes pour le diagnostic du TDC soulignent l'aspect multidisciplinaire de la pose du diagnostic de cette condition (ergothérapeute, psychomotricien, médecin, orthophoniste, neuropsychologue) et affirment l'importance de se référer à des outils standardisés qui limitent la subjectivité comme le MABC-2 et le DCDQ, tests les plus analysés. Il existe quantité d'outils diagnostics, une multitude de tests ou de questionnaires, il s'agira de définir lesquels sont efficaces.

1.1.3 Les outils diagnostics

La présence de facteurs de risques sont des indicateurs mais seuls les tests cliniques représentent des éléments de diagnostic, il s'agit donc de connaitre leur qualité, c'est-à-dire leur validité, leur fiabilité, leur reproductibilité et leur application potentielle en kinésithérapie.

On peut alors se demander pourquoi est-il important de poser le diagnostic, et cela, le plus tôt possible dans le développement de l'enfant ? La littérature s'accorde sur l'importance décisive d'une identification précoce mais est rendue complexe par l'absence d'un critère standard accepté et démontré pour les TDC donc une évolution multiple est utile, nécessaire et intéressante. Les études soulignent que, sans intervention, l'enfant est moins susceptible de s'améliorer dans ses difficultés motrices et plus l'intervention auprès de l'enfant est précoce, moins l'enfant est à risque d'accumuler les retards scolaires mais cela permet aussi une meilleure compréhension des besoins de l'enfant de la part de ses proches ainsi qu'une meilleure adaptation de la conduite des parents, des enseignants et des autres individus de son entourage. L'objectif étant de minimiser les répercussions émotionnelles et comportementales sur l'enfant : estime de soi, intimidation, isolement, sédentarité conduisant à une obésité, renoncement, etc. Il est de même nécessaire de faire comprendre que le TDC ne se guérit pas mais qu'il est indispensable d'apprendre à s'adapter et que l'effort devra se faire sur la longueur, si ce n'est toute sa vie.

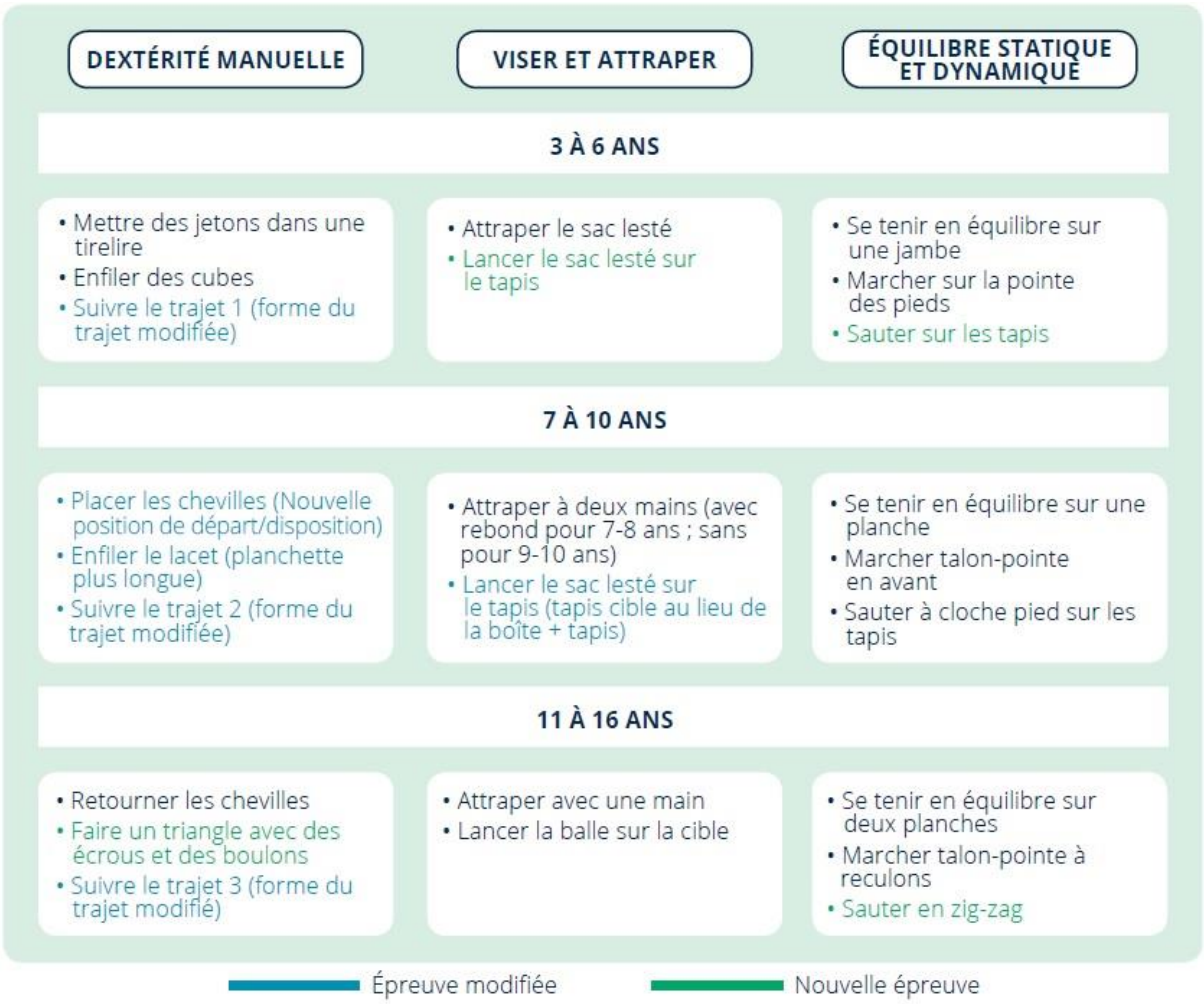
A ce titre, un MK peut avoir à prendre en charge un enfant TDC, pour une pathologie mécanique classique, telle qu'une entorse, une attitude scoliotique mais parfois le MK peut également faire partie de la prise en charge pluridisciplinaire, il est donc important qu'il s'intéresse aux outils de mesure parmi les plus analysés que sont le MABC-2 et le DCDQ.

1.1.3.1 Le MABC-2

Le MABC-2 (Movement Assessment Battery for Children - second edition) est une batterie d'évaluation du mouvement chez l'enfant de 3 ans à 16 ans 11 mois, il s'adapte à tous les âges. C'est le test de référence pour évaluer les troubles ou les retards du développement moteur. C'est une révision du MABC, trop long et compliqué à mettre en application, même si sa validité et sa fiabilité ont été démontrées, les professionnels ont préféré renouveler le modèle pour plus de simplicité et d'efficacité et pour s'adapter aux normes récentes. Elle se compose de deux parties, le test de performance et une liste à cocher. Le temps de passation du test est entre 20 à 40 min et le questionnaire environ 10 min. [10][11]

Ses points forts sont son extension de l'âge d'application, 3 à 16 ans 11 mois, les normes réactualisées et récentes et la rapidité d'administration et de cotation. Le MABC-2 fournit un aperçu précis des performances au cours des activités quotidiennes requérant de la coordination motrice. L'administration est individuelle et dure 20 à 40 minutes selon l'âge de l'enfant, ses potentialités et ses difficultés. La rapidité de l'épreuve réduit la fatigabilité, renforce la motivation et permet d'éviter de mettre l'enfant en échec. Trois groupes d'âges sont désormais proposés (de 3 à 6 ans ; de 7 à 10 ans ; de 11 à 16 ans 11 mois) : les épreuves, spécifiques à chaque groupe, sont adaptées au développement moteur qui le caractérise. Quatre nouvelles épreuves ont été créées et six épreuves ont été révisées pour proposer une évaluation réactualisée. Trois grands domaines sont explorés : la dextérité manuelle, les compétences viser et attraper puis l'équilibre statique et dynamique. [12]

Epreuves du test moteur :



En plus des données quantitatives provenant de l'examen formel, des descriptions qualitatives indiquant comment l'enfant réalise chacune des 8 épreuves peuvent être consignées par le professionnel. L'identification de ces caractéristiques comportementales perceptivo-motrices ainsi qu'un ensemble de facteurs non-moteurs susceptibles d'affecter la performance de l'enfant complètent l'examen. Le MABC-2 fournit des données quantitatives objectives relatives aux compétences psychomotrices : une note standard pour chacun des trois domaines est fournie ainsi qu'une note totale de test, reflétant la performance motrice globale de l'enfant. Un système d'alerte basé sur des seuils critiques (percentiles correspondants) indique si la note totale de test obtenue le situe en zone rouge « trouble moteur manifeste » (jusqu'à 5%), en zone orange « à risque » (5 à 15%) ou zone verte « performance dans la norme ». L'examen des résultats aux différents domaines permet d'établir un profil plus précis des difficultés de l'enfant.

Le questionnaire comprend désormais 43 items ciblés qui permettent d'explorer les difficultés et potentialités de l'enfant dans son quotidien (à la maison et à l'école) : l'exploration des capacités motrices selon le type d'environnement (Parties A et B : 30 items) ces parties permettent d'identifier les difficultés de l'enfant à réaliser certaines tâches dans des

conditions particulières : environnement statique et/ou prévisible, environnement dynamique et/ou non-prévisible. Puis les facteurs non-moteurs perturbateurs (Partie C : 13 items) cette partie permet de considérer dans quelle mesure des aspects du comportement de l'enfant peuvent entraver sa motricité : le manque d'attention, le manque d'organisation, etc. La liste de contrôle comprend ainsi 6 facteurs : la coordination motrice globale, les habiletés de balle, les compétences récréatives, la motricité fine, les habiletés rythmiques et l'équilibre dynamique. La note totale de motricité, obtenue à partir de la somme des notes des Parties A et B, permet de situer l'enfant dans un système d'alerte similaire au test (zones rouge, orange et verte) pour une grande simplicité d'interprétation.

Le MABC-2 est un test moteur étudié assez largement depuis plusieurs années pour confirmer ou infirmer ses qualités statistiques.

1.1.3.2 Le DCDQ

Le Questionnaire sur le Trouble Développemental de la Coordination (DCDQ) est un questionnaire parental permettant d'aider à l'identification du trouble développemental de la coordination (TDC) chez les enfants de 5 à 15 ans. Les parents doivent comparer le fonctionnement moteur de leur enfant à celui d'autres enfants du même âge selon une échelle de 5 points. Il est réalisable en une dizaine de minutes. Il est fondé sur le DCDQ'07, dont la cohérence interne est élevée et les résultats des analyses discriminantes sont appropriés pour un outil de dépistage. Le TDC est un diagnostic décrit dans le DSM-V. L'indication d'un risque de TDC selon le score obtenu au DCDQ remplit les conditions du critère B du diagnostic. Toutefois, le questionnaire ne peut être utilisé seul pour établir le diagnostic de TDC car celui-ci repose sur plusieurs observations et évaluations. Le DCDQ contient 15 questions regroupées en trois catégories distinctes. La première catégorie contient des questions reliées au contrôle moteur lorsque l'enfant ou un objet est en mouvement. Elle est intitulée « contrôle durant le mouvement ». La deuxième catégorie contient des questions concernant la « motricité fine et l'écriture » et la troisième catégorie contient des questions relatives à la « coordination globale ». Le score individuel de chacune de ces catégories ne fournit pas d'indication sur la présence ou non d'un TDC. Cependant, lorsque ces scores sont comparés entre eux et avec des résultats d'évaluations formelle et informelle, ils aident à l'identification des forces et/ou des difficultés motrices d'un enfant. Le DCDQ ne peut être utilisé seul pour identifier un TDC. Lorsqu'on se réfère aux résultats de ce questionnaire oralement ou dans un rapport écrit, les termes « indication d'un risque de TDC » ou « suspicion d'un TDC » ou « probablement pas de TDC » devraient être utilisés. Selon Wilson et Martini, pour les enfants âgées de 5 ans à 7 ans 11 mois, de 15 à 46 points il y a une indication d'un risque de TDC, de 47 à 75 probablement pas de TDC. Pour les enfants âgées de 8 ans à 9 ans 11 mois, de 15 à 55 points il y a une indication d'un risque de TDC, de 56 à 75 probablement pas de TDC. Pour les enfants âgées de 10 ans à 15 ans, de 15 à 57 points il y a une indication d'un risque de TDC, de 58 à 75 probablement pas de TDC. Le questionnaire est souvent proposé aux enseignants en complément, des personnes qui observent l'enfant dans des situations motrices du quotidien avec des résultats différents par rapport à ceux des parents. Le kinésithérapeute voyant un patient régulièrement, à l'occasion d'un protocole de rééducation, peut définir et constater des lacunes ou des difficultés motrices en association avec les parents. [13][14]

1.1.4 Traitement

1.1.4.1 Méthodes

Il existe plusieurs méthodes pour rééduquer les TDC et notamment trois : l'imagerie motrice, la CO-OP (Cognitive Orientation to daily Occupational Performance) et la NTT (Neuromoteur Task Training). On sait que les répétitions motrices et les démonstrations/modèles visuels n'aident pas à engrammer les mouvements et ne marchent pas avec les TDC.

Les axes généraux de prises en charge pour les TDC passent le plus souvent par le canal auditivo-verbal, c'est-à-dire la parole et on utilise les aspects cognitifs également.

La prise en charge du TDC porte à la fois sur des aspects relationnels et de motivations auprès de l'enfant, de ses parents mais aussi du milieu scolaire. Il y a 2 types de prise en charge, les approches centrées sur le déficit et sur le processus (bottom-up = Bas en haut, activation répétée va activer la fonction (plasticité), entrainement intensif et répétitif de la fonction lésée et rétablissement de la fonction sous tendant l'activité) avec la thérapie d'intégration sensorielle et kinesthésique puis les approches centrées sur la performance et sur la tâche (top down = Haut en bas, activation des fonctions supérieures va activer les fonctions inférieurs (plasticité), rééducation cognitive ou métacognitive et prise de conscience des déficits et auto-régulation) avec la NTT, la CO-OP et l'imagerie motrice qui sont les méthodes préconisées même si l'imagerie motrice est encore à l'étude. [15]

1.1.4.2 Imagerie motrice

L'imagerie motrice s'appuie sur l'utilisation des neurones miroirs, c'est une technique déjà utilisée dans d'autres domaines notamment le sport de haut niveau. Cette méthode est un processus de perception en l'absence de tout stimulus externe et s'appuie sur la capacité de visualiser des mouvements et des actions faisant référence à une image que le participant expérimente comme s'il pratiquait l'action volontairement (souvent d'abord à la troisième personne puis idéalement à la première). Les images dynamiques des actes moteurs se concentrent principalement sur les informations kinesthésiques ou visuelles d'une perspective à la première personne. Le protocole est le suivant : Exercice de visualisation et de timing, protocole de relaxation et de préparation mentale, visionnage de modèles d'habilités motrices fondamentales, exercices de visualisation des habilités motrices depuis une perspective externe, des exercices de visualisation motrice depuis une perspective interne et enfin la pratique ouverte (l'action replay par exemple). Cela facilite la planification prédictive des paramètres du mouvement tels que la force, le timing, et la trajectoire Mode du protocole. La méthode encourage la personne qui exécute le mouvement à prévoir les conséquences pour l'action en l'absence de mouvement visible.

Par un processus de répétition mentale, l'enfant « visualise » le flux temporel de l'espace visuel qui accompagne le mouvement normal et, en adoptant une perspective à la première personne », ressent « comment » ce mouvement est calibré à travers les sensations kinesthésiques au niveau des membres. Au fil du temps et avec la pratique, les enfants utilisent la connaissance de cette relation entre la vision et la kinesthésie pour faire des prédictions précises sur les conséquences du mouvement. Ainsi, les erreurs dans la planification d'action directe sont atténuées. L'amélioration, en fin de compte, consiste en la

capacité de planifier et de coordonner les membres lors du mouvement par rapport aux objets et aux obstacles de l'environnement. Les objectifs de progression sont le timing prédictif, la relaxation et la préparation mentale et enfin le modelage et l'imagerie mentale.

1.1.4.3 NTT (Neuromoteur Task Training)

La NTT est un programme de traitement orientée sur la tâche basé sur les théories sur le contrôle et l'apprentissage moteur. Le thérapeute utilise 3 catégories de stratégie : donner des instructions, apporter des feed-backs ou en demander à l'enfant et partager des connaissances. C'est en fait une technique à partir des théories d'apprentissage moteur, elle se focalise sur la structure de la tâche, sur la façon dont les compétences sont assemblées au cours des séances et sur des essais répétés d'apprentissage. Le thérapeute doit objectiver le ou les processus de contrôle moteur déficitaire chez l'enfant. Pour cela, le thérapeute effectue une évaluation minutieuse lors de l'exécution d'une habilité motrice fonctionnelle. Pour les théories écologiques, la NTT considère la façon dont les contraintes de travail et de l'environnement peuvent être manipulées pour fournir un effet de levier pour l'individu dont la difficulté motrice ne permet pas directement un mouvement plus fluide dans un espace particulier. La NTT met l'accent sur les principes pédagogiques utilisés pour guider la pratique et l'importance de fournir une rétroaction pour améliorer l'apprentissage. Les enfants TDC sont moins performants pour sélectionner la bonne réponse motrice (répétition des erreurs malgré les échecs). En conséquence, la répétition et les essais spontanés ne vont pas leur permettre d'améliorer leurs performances. Il y a d'abord analyse de l'exécution de l'habilité par le professionnel puis un travail spécifique du processus déficitaire avec une variation des contextes d'apprentissage. Enfin, ce qui est déterminant, c'est la manière dont le thérapeute va donner les instructions mais aussi apporter des feed-back à l'enfant ou lui en demander. En effet, c'est la manière de donner de l'information et des sensations sensorimotrices (retour proprioceptif par exemple) qui va être déterminante durant les exercices. La NTT décompose ainsi le contrôle neuromoteur avec les 6 processus de contrôle impliqués :

- Perception : l'apprentissage moteur permet ainsi d'acquérir de l'expérience dans des circonstances variées en améliorant les processus de prise d'information
- Planification de l'action : apprendre à faire des choix selon les changements de condition stimulus/réponse (si/alors)
- Planification du mouvement : améliorer l'exécution de l'habilité avec une augmentation de la complexité et de la durée de la séquence
- Programmation motrice : l'apprentissage de certains patterns et séquences, essai de déviation du mouvement
- Paramétrisation : l'entrainement aux variations en ajustant l'extension, la vitesse et la direction du mouvement, l'apprentissage de comment utiliser les caractéristiques biomécaniques
- Force initiale : l'entrainement aux réglages fins en faisant des choix (type/nombre) des unités motrices dans différents contextes.

1.1.4.4 CO-OP (Cognitive Orientation to daily Occupational Performance)

Enfin la CO-OP est une méthode qui a montré son efficacité sur les tâches ciblées au niveau de la généralisation et le transfert des habilités. Elle a un mode d'action motivationnel, ce sont

les enfants qui choisissent les habilités à travailler (motivation). Le thérapeute utilise la stratégie globale de manière systématique et utilise, selon les difficultés et les ressources de l'enfant, une à plusieurs stratégies spécifiques pour permettre l'apprentissage. Les théories sous-jacentes sont actuelles et comprennent à la fois les théories cognitives et les théories dynamiques. C'est une méthode dite cognitive qui associe une technique de résolution de problèmes moteurs avec l'utilisation du guidage verbal. Le thérapeute permet la découverte des stratégies efficaces mais ne les donnes pas directement. Le TDC est envisagé par la méthode CO-OP comme un trouble touchant la capacité à apprendre des coordinations motrices. La remédiation doit alors s'envisager comme une technique permettant l'acquisition d'habiletés dans une perspective d'apprentissage. Il y a sept points clés : objectifs choisis par les patients, analyse dynamique de l'exécution, utilisation de stratégies cognitives, découverte guidée, principes facilitants, implication parents/entourage et le format intervention. Il existe 4 principes facilitants : rendre les choses amusantes, favoriser l'apprentissage, favoriser l'autonomie et promouvoir la généralisation et le transfert. Il s'agit de pratiquer la découverte guidée avec une seule chose à la fois, poser des questions, guider, ne pas faire à la place et rendre les choses évidentes. Deux types de stratégies sont connus, la stratégie globale qui est une technique simplifiée de résolution de problème (goalplan-do-check) et les stratégies spécifiques qui vont permettre de résoudre des difficultés particulières de l'enfant sur la tâche en cours. Durant la ou les premières séances CO-OP, il y a une phase cognitive d'apprentissage de cette stratégie, c'est-à-dire que l'on va apprendre les étapes en dehors de toute pratique concrète. Par la suite, on demande aux enfants de l'utiliser à chaque passage effectif sur la tâche choisie. Les stratégies spécifiques s'appuient sur 7 stratégies : la position du corps, l'auto-guidance verbale, se focaliser sur les points d'exécution, image motrice mnémonique, sentir le mouvement et la modification ou aménagement de la tâche. Le thérapeute va également juger la progression avec un arbre décisionnel échec/réussite.

Ces traitements sont les fondamentaux de la rééducation psychomotrice en particulier. Les TDC sont traités selon leur profil, chaque enfant est différent et l'éventail offert permet d'aider les patients.

1.1.5 Intérêts de cette revue de littérature

Les premiers travaux en France datent de 1964, ce qui nous montre à quel point, la recherche est récente et la pathologie mal connue dans la société civile mais aussi des professionnels de santé. Le bref retour historique sur les terminologies utilisées dans la littérature scientifique indique que le TDC fait l'objet d'un champ de recherche récent dont le périmètre s'affine progressivement. S'il faut reconnaître des avancées avec une certaine harmonisation permise par l'usage du terme TDC associé, plus ou moins rigoureusement, aux critères du DSM, les publications sur ce thème restent de qualité très inégale et appellent donc à une grande vigilance dans l'interprétation des résultats. Par ailleurs, dans la dynamique de développement du champ, on voit aussi que certaines disciplines ou thématiques d'intérêt sont encore très peu explorées et notamment la kinésithérapie qui est complètement absente du parcours de soin alors qu'elle est pourtant, par essence, la discipline du mouvement perturbé. Le mot kinésithérapie, tire son origine du grec kínesis (« mouvement »), il est donc habilité à intervenir dans le traitement des troubles moteurs, ici les TDC [16]. De plus le

kinésithérapeute est un professionnel de santé qui côtoie l'enfant tout au long de son développement dans des circonstances thérapeutiques variées permettant de noter l'évolution motrice du jeune et de constater des difficultés motrices potentielles, dans la prise en charge de problèmes parfois uniquement mécaniques (entorse, scoliose, déchirure musculaire), avec ou sans lien avec les TDC, ou dans la prise en charge pluridisciplinaire des TDC.

Cette revue présente plusieurs intérêts. Tout d'abord, les TDC sont désormais un problème de santé publique et diffuser de la littérature scientifique sur le sujet est toujours utile. Il est important de concerner l'ensemble du monde médical sur un sujet qui est méconnu ou mal compris. Les TDC sont souvent sous diagnostiqués et la souffrance que peut ressentir l'enfant dyspraxique doit nous pousser à resserrer les mailles du diagnostic et renforcer la vigilance des professionnels de santé, les kinésithérapeutes compris. Il faut impérativement que le monde de la kinésithérapie se sente concerné et intègre le domaine de la rééducation des TDC. Le kinésithérapeute, en utilisant la thérapie par le mouvement, a la chance de voir son patient dans des situations de rééducation qui peuvent le mener à des échecs répétés avec des exercices qui peuvent lui être complexes alors qu'ils le sont beaucoup moins en temps normal et il doit alors être capable de participer au parcours de soin en identifiant une difficulté anormale ou pathologique ressemblant à un TDC. Il ne faut pas cloisonner le traitement des TDC et il s'agit ainsi de participer à la détection d'un trouble et l'identifier le plus clairement possible puis l'orienter vers un médecin ou un spécialiste. Pour définir si le patient est un TDC, des tests cliniques accessibles sont susceptibles d'être utilisés à ce stade, des questionnaires destinés à l'entourage et au patient lui-même ou encore interroger sur les potentiels facteurs de risques. Si le patient répond positivement aux critères conduisant à un TDC, le kinésithérapeute se doit alors de s'adapter à son patient et modifier son traitement et sa rééducation en fonction des défauts identifiés. Le kinésithérapeute, même s'il ne remplacera jamais un psychomotricien, se doit de se renseigner le cas échéant, et connaître des principes de base de la prise en charge du TDC. Il y a un intérêt pour le kinésithérapeute qui va avoir un meilleur résultat pour son programme de rééducation mais aussi un bénéfice important pour l'enfant. Il peut se positionner en complément, en soutien et jouer un rôle important voir décisif dans l'orientation du patient notamment mais aussi la compréhension, la considération et la prise en charge adéquate. Le TDC est souvent mal aiguillé ou pas considéré comme malade, qu'il n'y a pas de progression possible, il s'adaptera tout seul. Je suis moi-même un ancien enfant TDC détecté à 9 ans par hasard à l'occasion d'un autre examen mais non orienté vers un professionnel et les réelles difficultés se sont montrées à mon arrivée à l'école de kinésithérapie et c'est seulement aujourd'hui que je suis traité, je comprends donc la nécessité voire l'urgence de la rééducation et du rôle que le kinésithérapeute pourrait prendre. Le principal est d'agir dans l'intérêt des patients et que le kinésithérapeute puisse s'enrichir de compétences pour apporter une solution supplémentaire au patient TDC pour le profit du patient mais aussi du kinésithérapeute.

De plus, le coût de la santé est un sujet de société important et épineux aujourd'hui. Il s'agit de conserver la meilleure qualité de soins possibles tout en limitant les dépenses inutiles. Le rapport qualité prix est une donnée à prendre en compte, le DCDQ est peu coûteux et facile à mettre en place alors que le MABC-2 est plus pratique que son prédécesseur moins long et

sans matériel spécifique, ce sont des outils diagnostics qui semblent accessibles et universels. Il ne faut malgré tout pas négliger la validité et la fiabilité de l'instrument ainsi nous allons donc étudier la littérature pour savoir si leur utilisation est efficace.

1.1.6 Objectifs de cette revue de littérature

Modèle PICO	Explication
Population	Enfants de 5 à 16 ans
Intervention	Outils de mesure (DCDQ, MABC-2)
Comparateur	
Outcome (critère de jugement)	Propriétés métrologiques (validité et fiabilité)

Tableau 1 : descriptif de la revue selon le modèle PICO

On peut ainsi se demander comment le kinésithérapeute peut participer au diagnostic de TDC de patients atteints et ainsi comment en tant qu'acteur de « la santé du mouvement perturbé », nous devons être à même de dépister cette pathologie parmi nos patients afin de les traiter avec efficacité. Pour développer ce domaine de compétences, nous devons commencer par nous questionner sur les outils d'évaluation existants, leur validité et leur fiabilité, leur utilisation adaptée aux kinésithérapeutes, afin de faire un choix éclairé pour notre pratique.

Le MABC-2 et le DCDQ, outils diagnostics du TDC déjà existants, permettent-ils d'identifier efficacement les TDC ?

2 Méthodologie

Afin de réaliser une revue systématique de littérature dans un temps imparti, il a été nécessaire d'élaborer une méthodologie de recherche pertinente, structurée et efficace. Cette partie portera sur ma stratégie de recherche, c'est-à-dire la méthodologie employée pour construire mes équations de recherche et sélectionner mes études, ainsi que sur la méthode d'évaluation des études sélectionnées me permettant d'aboutir à une synthèse des résultats.

2.1 Critères d'éligibilité des études pour cette revue

2.1.1 Types d'études

Le type d'études le plus apte à répondre à la question clinique est l'étude transversale [17]. Selon l'HAS dans l'état des lieux sur le niveau de preuve et gradation des recommandations de bonne pratique de 2013, le type d'étude adéquate pour chercher à évaluer la reproductibilité est l'étude transversale comparative avec répétition de mesure alors que pour juger la sensibilité/spécificité, ce sera l'étude transversale comparative avec étalon-or. C'est une étude évaluative où la qualité d'un test diagnostic (validité et fiabilité) sera estimée. [18][19]

Grade des recommandations	Niveau de preuve scientifique fourni par la littérature
A Preuve scientifique établie	Niveau 1 - essais comparatifs randomisés de forte puissance ; - méta-analyse d'essais comparatifs randomisés ; - analyse de décision fondée sur des études bien menées.
B Présomption scientifique	Niveau 2 - essais comparatifs randomisés de faible puissance ; - études comparatives non randomisées bien menées ; - études de cohortes.
C Faible niveau de preuve scientifique	Niveau 3 - études cas-témoins.
	Niveau 4 - études comparatives comportant des biais importants ; - études rétrospectives ; - séries de cas ; - études épidémiologiques descriptives (transversale, longitudinale).

Tableau 2 : "État des lieux : Le niveau de preuve et gradation des recommandations de bonne pratique" (HAS, 2013)

Question	Protocole
THÉRAPEUTIQUE Efficacité	Étude contrôlée randomisée
THÉRAPEUTIQUE Sécurité	Étude contrôlée randomisée ou suivi de cohorte
DIAGNOSTIC Reproductibilité/Variabilité	Transversal comparatif avec répétition de mesure
DIAGNOSTIC Sensibilité/Spécificité	Transversal comparatif avec étalon-or
DIAGNOSTIC Efficacité/Utilité	Étude contrôlée randomisée
DIAGNOSTIC Stratégie	Étude contrôlée randomisée ou arbre décisionnel
CAUSALITÉ Phénomène contrôlable fréquent	Étude contrôlée randomisée
CAUSALITÉ Phénomène non contrôlable fréquent	Suivi de cohorte (exposés/non exposés)
CAUSALITÉ Phénomène rare	Étude cas-témoin
PRONOSTIC Maladie fréquente	Étude contrôlée randomisée ou suivi de cohorte
PRONOSTIC Maladie rare	Étude cas-témoin

Tableau 3 : Type de protocole préférentiellement proposé pour une question donnée extrait du document « État des lieux : Le niveau de preuve et gradation des recommandations de bonne pratique" (HAS, 2013) »

2.1.2 Population/Pathologie

Dans une étude transversale, l'échantillon est issu d'une population sans être sélectionné sur l'exposition (cohorte) ni sur la maladie (cas témoins).

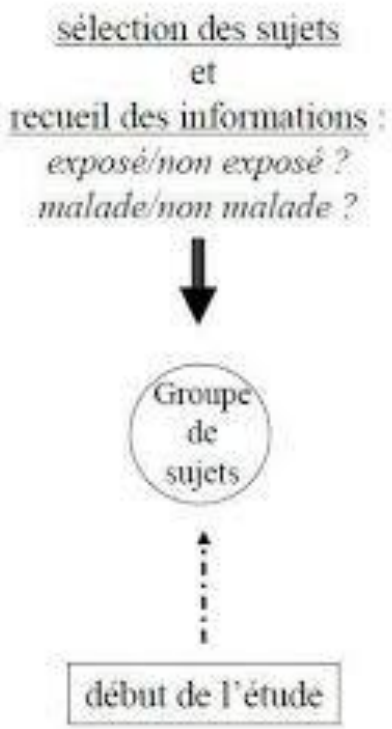


Tableau 4 : Schéma général d'une étude transversale

Ce mémoire porte sur l'évaluation d'un test diagnostic de TDC qui s'adresse à des enfants de 5 à 15 ans pour le DCDQ et de 3 à 16 ans pour le MABC-2. La population cible est ainsi un échantillon d'enfants dans cette tranche d'âge ou une sous-catégorie d'âge de l'outil mesuré (3 sous catégories distinctes présentées précédemment dans chacun des deux outils). Les études s'adressent généralement à une population avec et sans troubles moteurs pour juger

les capacités d'identification des TDC dans un groupe représentatif de la classe d'âge mais peut aussi s'adresser à un échantillon d'enfants présentant des difficultés motrices pour juger de sa capacité à déterminer les TDC dans une population à risque et potentiellement proche de la population qui va être testée dans la pratique clinique. Les enfants avec des troubles neurologiques associés sont exclus, notamment le spectre autistique, les TDC étant très souvent présents dans les articles sur ce sujet.

2.1.3 Outil de mesure

L'étude concerne ainsi le dépistage de TDC. On procède ainsi à l'évaluation de l'échantillon avec un élément de référence, un gold standard le plus étudié dans la littérature scientifique et validé possible, pour avoir l'estimation référente et ainsi comparer les résultats du DCDQ ou du MABC-2 à ce dernier. Le DCDQ est un questionnaire à faible coût, facilement compréhensible adressé aux parents alors que le MABC-2 est un test clinique lié à une liste de contrôle réalisé par un professionnel.

2.1.4 Objectif/critères de jugement

L'intérêt de ce mémoire est ainsi de mesurer la validité prédictive (sensibilité et spécificité, la valeur prédictive positive et négative) [20][21], la validité de construction (validité discriminatoire, validité convergente et validité de groupe connue), la validité concomitante (compare à gold standard), la validité inter culturelle mais aussi la fiabilité, c'est-à-dire la cohérence globale d'une mesure, si elle est stable lorsqu'elle est répétée dans des conditions cohérentes, la fiabilité test-retest donc stable dans le temps ou encore la fiabilité inter-évaluateur (l'homogénéité) et la cohérence interne, c'est-à-dire les éléments du questionnaire mesurent la même construction (coefficient α de Cronbach).

2.2 Méthodologie de recherche des études

2.2.1 Sources documentaires investiguées

L'objectif étant de restituer un rapport sourcé de la qualité du dépistage des TDC, il a été nécessaire d'établir une méthode de recherche fiable, pertinente et efficace en raison du temps de travail et des ressources limités sur ce sujet dans le domaine de compétence de la kinésithérapie. Un travail d'investigation sur les diverses sources d'informations officielles a été nécessaire pour permettre de traiter les points clés majeurs de la problématique. J'ai effectué mes recherches sur des livres de psychométrie au sein des bibliothèques universitaires, sur le site de l'Inserm et notamment les recommandations de bonnes pratiques publiées par cet institut pour me documenter sur le sujet, des revues scientifiques, Google books ainsi que le site de l'HAS. Mon travail est essentiellement basé sur les moteurs de recherche suivants : PubMed, Cochrane, Embase, Science direct, Google Scholar et Kinedoc. La bibliothèque universitaire en ligne m'a permis d'accéder à de nombreux articles. Des recherches dans la littérature grise incluant des mémoires, des thèses, ont également été effectuées.

2.2.2 Equation de recherche utilisée (mots clefs)

Pour chaque base de données scientifiques, il a été établi une équation de recherche spécifique structurée avec des mots-clés et des termes Medical Subjects Heading (MeSH) en lien avec mon sujet à l'aide d'opérateurs booléens (OU/OR ; ET/AND ; PAS/NOT). Toutefois, avant de me lancer dans la construction de mon équation de recherche finale, j'ai tout d'abord commencé par une recherche simple sur les différentes bases de données, afin d'avoir une vue d'ensemble sur la bibliographie en lien avec mon sujet, puis j'ai affiné mes requêtes à l'aide de la recherche avancée dans le but d'obtenir des articles pertinents. J'ai ainsi utilisé une première équation basique en cherchant avec seulement « DCD OR dyspraxia AND diagnostic NOT autism » pour me faire une première impression de la littérature en présence. J'ai ensuite amélioré ma recherche en me concentrant sur le DCDQ et le MABC-2. La recherche se fait en anglais et avec des termes MeSH. Afin d'identifier les mots clés, les synonymes et les termes MeSH en anglais, associés aux termes de mon sujet, j'ai utilisé le thésaurus MeSH dans les bases de données PubMed. Grâce à l'outil de recherche avancée, nécessitant la construction d'une équation de recherche structurée via des opérateurs booléens (AND, OR, NOT), j'ai pu obtenir des publications en adéquation avec mon sujet.

L'équation de recherche définitive est « DCDQ OR MABC-2 AND reliability AND validity ». Le DCDQ et le MABC-2 étant des outils s'adressant à des enfants sans troubles autistiques, il n'était pas nécessaire de conserver l'exclusion dans l'équation. Je cherche la validité et la fiabilité de ces derniers, il reste ainsi 36 articles sur PubMed et 143 sur la bibliothèque universitaire en ligne avec des doublons.

La simplicité de cette équation de recherche a été un choix, au vu des résultats obtenus, dans un but de ne pas être trop restrictif et éliminer des articles susceptibles d'être intéressant pour notre questionnement. Notre but étant ici de regrouper tout ce qu'il se fait ici concernant le diagnostic des TDC grâce au DCDQ et le MABC-2

2.3 Extraction et analyse des données

2.3.1 Sélection des articles

L'équation de recherche finalement validée, les articles sont sélectionnés en premier lieu par le titre et le résumé en fonction des critères d'éligibilité préalablement établis. Si aucune information n'était contenue en lien avec mes critères, ces derniers étaient exclus. Il faut ainsi qu'apparaissent les mots clés suivants : DCDQ ou MABC-2 de façon indispensable mais aussi DCD, Motor Skills Disorders, Child Development, Validation, Validity, Specificity, Sensibility, Reliability. Tous les doublons ont pu de cette façon être identifiés et supprimés des articles potentiellement inclus. Parmi tous les articles restants, ceux dont la pertinence me semblait en adéquation avec ma problématique étaient sélectionnés pour qu'ils soient lus en entier. La deuxième étape consiste alors en la lecture complète de ces articles en vérifiant si les critères d'éligibilité étaient respectés ainsi que la qualité des études. A partir des bibliographies des articles, j'ai pu être redirigé vers de nouveaux articles intéressants que j'ai également réintégré dans la sélection pour les évaluer et savoir au même titre que les autres s'il fallait les intégrer à la revue.

2.3.2 Extraction des données

Pour lire entièrement les articles sélectionnés, ceux-ci ont été récupérés soit en téléchargement libre, soit en utilisant la bibliothèque universitaire pour avoir un accès à l'article grâce à Aix-Marseille Université.

Une première lecture complète des articles a été réalisée en rapport aux critères d'éligibilité et une seconde lecture pour extraire les différentes données utiles. Les articles retenus ont été analysés, et les informations suivantes ont été relevées : le nom des auteurs, l'année de publication de l'article, le type d'étude, la population étudiée, la taille de l'échantillon, les détails de l'étude, l'outil de dépistage et le comparateur utilisé mais aussi les praticiens, la procédure de retest, les propriétés métrologiques de l'étude ainsi que les paramètres statistiques. Les différents articles retenus en lecture intégrale ont été répartis sous la forme d'un tableau synthétique. Ce type d'illustration permet au lecteur de visualiser les informations essentielles de chaque article, et une synthèse plus détaillée de chaque article se trouvera en annexe.

2.3.3 Evaluation de la qualité méthodologique des études sélectionnées

Les articles retenus par lecture complète ont été évalués par les échelles de qualité méthodologique QUADAS-2 et QAREL qui définissent la qualité des études diagnostiques à inclure dans les revues systématiques en étudiant la validité pour l'une, la fiabilité pour l'autre et se rapportent au rapport de recommandations STARD, grille qui définit les éléments devant figurer dans un article rapportant les résultats d'une étude diagnostique.[22]

L'étude diagnostique de validité est une étude transversale qui va analyser l'association entre des mesures d'un test et d'une entité comparatrice (référence standard). L'échelle de qualité méthodologique pour coter la validité interne de ce type d'étude est l'échelle QUADAS-2. Elle est conçue pour évaluer la qualité des études de précision diagnostique primaire et pas pour remplacer le processus d'extraction des données de l'examen et doit être appliquée en plus de l'extraction des données primaires. La QUADAS-2 est répartie en trois phases : la première au sujet de la question clinique c'est-à-dire les patients (le contexte, le motif de consultation, le tableau clinique), le test évalué et le test de référence ; la deuxième avec le diagramme de flux primaire de l'étude ; la troisième concerne le risque de biais et le jugement sur l'applicabilité clinique de la question de recherche. Cette dernière étape de la QUADAS-2 est organisée en quatre domaines et chaque domaine comprend une série de questions qui permettent de porter un jugement sur le risque de biais et l'applicabilité clinique. L'outil se compose de ces quatre domaines clés qui discutent de la sélection des patients, du test évalué, de l'étalon de référence mais également du flux de patients tout au long de l'étude et du calendrier du test évalué et de l'étalon de référence (flux et timing) [23][24]. L'étude diagnostique de fiabilité est une étude transversale qui va analyser l'association entre des mesures répétées à partir d'un test (réalisées par un ou plusieurs évaluateurs). L'échelle de qualité méthodologique pour coter la validité interne de ce type d'étude est l'échelle QAREL. Cette dernière comporte onze items concernant l'échantillon, l'évaluateur, l'ordre d'évaluation, la stabilité de la variable, la qualité de l'application de l'examen et l'utilisation des mesures statistiques appropriées.[25][26]

La grille STARD décrit les participants, les méthodes des tests, les résultats des tests, l'estimation et la discussion.

2.3.4 Méthode de synthèse des résultats

Une fois l'analyse sur l'échelle QUADAS-2 réalisée, les données des articles inclus ont été récupérées puis analysées et interprétées. Les données statistiques essentielles ont été ensuite organisées sous forme de tableaux. La rédaction de synthèse des résultats se fera par un tableau récapitulatif pour chacune des études incluses regroupant plusieurs critères : le design de l'étude ; la population (cible, effectif, mode de recrutement, …) ; l'intervention c'est-à-dire le mode opératoire de l'étude (l'outil mesuré, le comparateur gold standard, la procédure d'administration, etc.) ; les critères de jugement (validité et fiabilité) ; les résultats et les effets de l'intervention. Les résultats statistiques des études pourront ainsi être discutés en extrayant les informations déterminantes de ce tableau. Nous dresserons également un tableau récapitulatif des études exclues. Chaque étude sera comparée en fonction de mêmes critères prédéfinis. La synthèse des résultats a été traitée de manière quantitative, par groupe de comparaison.

3 Résultats

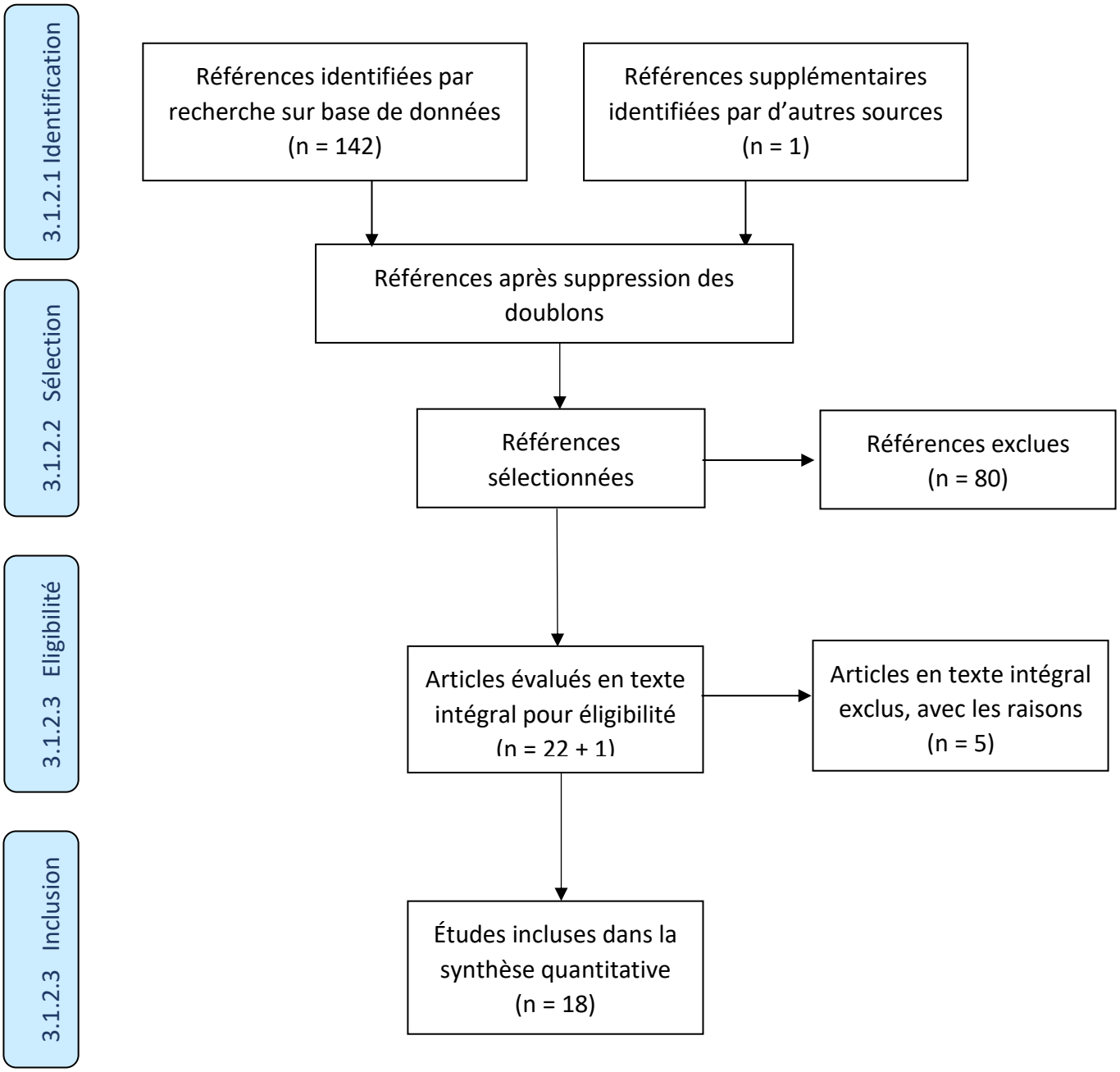
3.1 Description des études

3.1.1 Résultats de la recherche bibliographique - sélection des études

Sur les dix dernières années, la stratégie de recherche a fourni 143 études sur la bibliothèque universitaire en ligne mais aussi 36 études sur PubMed, 80 sur Google Scholar et la recherche sur PEDro avec l'équation de recherche ou sur Cochrane Library s'est avérée moins fructueuse. La bibliothèque universitaire m'a permis d'accéder à Science Direct ou encore d'autres bases de données scientifiques. Un psychomotricien m'a également transmis une étude française. Parmi les 143 articles, après suppression des doublons nombreux, la liste d'articles s'est réduite à 107, la première étape de sélection se concentre ainsi sur le titre et le résumé. Sur la totalité des études trouvées, 22 articles ont été sélectionnés par le titre et le résumé en excluant les études concernant les enfants avec des troubles neurologiques ou l'autisme, les études travaillant sur le MABC, première version du MABC-2 mais aussi les articles en langue étrangère hors de l'anglais et du français.

Ainsi, 22 essais potentiellement éligibles ont été examinés en vue de leur éligibilité par lecture intégrale. A la suite de la lecture complète, 5 articles ont été exclus par non-respect des critères d'inclusion et d'exclusion préalablement cités pour différentes raisons exposées plus tard. Un article a été inclus dans la revue à la suite de la lecture de leur bibliographie pour son travail sur la fiabilité inter-évaluateurs, rarement étudiée dans la littérature scientifique, et de nombreux autres ont été utiles pour la documentation. Il en résulte 18 articles respectant les critères d'inclusion. Par la suite, ils ont été évalués avec les échelles QUADAS-2 et QAREL. Au total : 18 études ont été incluses dans l'analyse finale pour constituer cette revue systématique.

3.1.2 Diagramme de flux



3.1.3 Caractéristiques des études exclues

Sur la quantité d'articles sélectionnés dans un premier temps, 85 ont été exclus à la lecture des titres ou des résumés. Les articles dans une langue étrangère différente de l'anglais ont été exclus au même titre que les articles sur le MABC qui étaient par confusion sémantique régulièrement dans cette sélection initiale, les articles où la population cible est composée d'enfants avec des troubles neurologiques ou bien les études portaient sur les TDC dans la

manifestation clinique de l'autisme. Certains autres étaient trop précises concernant les outils diagnostics soit l'objectif de l'étude n'était pas clair ou encore elles étaient trop évasives par rapport à mon sujet. Certaines étaient malheureusement indisponibles à la lecture complète.

Les études exclues par lecture intégrale n'ont pas été nombreuses. Cinq études ont été exclues par non-respect des critères d'inclusion pour les raisons suivantes :

Etudes	Raisons de l'exclusion
Sarraff T, Martinez C, Santos J Revista de Terapia Ocupacional da Universidade de São Paulo, 2018 [27]	Point de coupure pour chaque âge spécifique pas clairement défini, courbe ROC incohérente et non interprétable, trop petit effectif pour la classe de 9 ans
Dido Green, Tracie Bishop, Brenda N Wilson, Susan Crawford, Richard Hooper, Bonnie Kaplan and Gillian Baird British Journal of Occupational Therapy 2005 [28]	Non exclusion des enfants avec des troubles envahissant du développement, un spectre autistique, des troubles attentionnels (déficit de l'attention avec ou sans hyperactivité)
Kokštejn J, Musálek M, Tufano J Frontiers in Pediatrics 2018 [29]	Objectifs de l'étude ne correspondent pas aux critères de jugement de ma revue
Psotta R, Brom O Perceptual and Motor Skills 2016 [30]	Analyse factorielle et sous factorielle de structure uniquement, différence avec le critère de jugement de l'étude
Lauren R. Civetta, Susan L. Hillier Pediatric Physical Therapy 2008 [31]	Seuls les patients ayant eu une évaluation suspecte au DCDQ ont reçu le test moteur (gold standard) donc il y a une non-évaluation d'une grande partie des enfants avec un risque de faux positifs et faux négatifs

3.1.4 Caractéristiques des études incluses

Design : Dix-sept articles ont été inclus dans cette revue par respect des critères d'éligibilité et à la suite de l'évaluation de la qualité méthodologique avec l'échelle QUADAS-2 présentant un score de 4/7 au minimum pour les études objectivant la validité des outils diagnostics et avec l'échelle QAREL présentant un score de 8/11 pour les études objectivant la fiabilité des outils diagnostics. Ce sont toutes des études transversales comparatives ou non, à l'exception d'une revue systématique.

Population : La population cible est une population d'enfants de 3 à 16 ans pour le MABC-2 ou de 5 à 15 ans pour le DCDQ, soit sur l'entièreté de la classe d'âge, soit une sous population représentant un sous-groupe des outils évalués. Les échantillons sont, la plupart du temps, composés d'enfants non diagnostiqués et non suspectés la plupart du temps mais parfois aussi avec un groupe témoin et un groupe clinique. Ces échantillons sont caractéristiques de la population en étant originaire d'horizons socioéconomiques différents, avec des filles et des garçons et en excluant de l'effectif les enfants avec des troubles moteurs/physiques, intellectuels/mentaux ou neurologiques.

Intervention : Les outils de mesure qui sont évalués sont le DCDQ et le MABC-2. L'objectif de ce mémoire est de déterminer la validité et la fiabilité de ces derniers grâce par exemple à des études d'adaptations interculturelles, à des études s'intéressant aux propriétés métrologiques ou psychométriques des outils évalués et en se comparant à des références « gold standard ».

Critère de jugement : Ce travail s'intéresse à plusieurs critères qui permettent d'objectiver les qualités d'une étude. Il va ainsi mesurer la validité prédictive, la validité de construction, la validité concomitante, la validité inter culturelle mais aussi la fiabilité, c'est-à-dire la cohérence globale d'une mesure, si elle est stable lorsqu'elle est répétée dans des conditions cohérentes, la fiabilité test-retest donc stable dans le temps ou encore la fiabilité inter-évaluateur (l'homogénéité) et la cohérence interne (les éléments du questionnaire mesurent la même construction). L'intérêt est donc de définir si le MABC-2 ou le DCDQ sont efficaces dans son objectif de diagnostic des TDC.

Résultats : Ce sont les propriétés métrologiques qui sont retirés de ces études : la sensibilité, la spécificité, la valeur prédictive positive et la valeur prédictive négative, la courbe ROC, la cohérence interne avec l'α de Cronbach, la fiabilité test-retest ou inter-évaluateurs, le coefficient de corrélation de Pearson r ou celui de Spearman ρ, selon que la relation est linéaire ou monotone de type non affine, qui définit de quelle façon deux variables sont associées, la valeur de probabilité p qui informe si c'est significatif statistiquement.

3.1.5 Tableau synthétique

Auteurs/Année	Type de l'étude	Taille de l'échantillon	Population étudiée	Outil mesuré	Comparateur(s)	Echelle(s) Méthodologique(s)
Ferreira L et al, 2019 [32]	Etude Transversale Comparatif	n=707 (332G/375F) Groupe 1 n=206 Groupe 2 n=352 Groupe 3 n=149	6-7 ans et 11 mois = Groupe1 8-9 ans et 11 mois = Groupe 2 10-10 ans et 11 mois = Groupe 3 Pas d'handicap mental et moteur, statut socioéconomique différent	DCDQ	MABC-2/BOT-2	QUADAS-2
Martini R et al, 2011 [33]	Etude Transversale Comparatif	n=84	Sujets de 5 à 15 ans	DCDQ	MABC-2	QUADAS-2 QAREL
Montoro A et al, 2016 [34]	Etude Transversale Comparatif	n=350 (175F/175G)	Enfants de 7 à 10 ans et 11 mois Niveau socioéconomique différent Pas d'handicap moteur et mental	MABC-2	DCDQ	QUADAS-2
Capistrano R et al, 2015 [35]	Etude Transversale Comparatif	n=40 (20F/20G)	Enfants de 7 à 10 ans et 11 mois Pas d'handicap moteur et mental	DCDQ	MABC-2 Test et MABC-2 Checklist	QUADAS-2

Schoemaker M et al, 2012 [36]	Etude Transversale Comparatif	n=383 (190G/193F) → n=130 ont reçu le gold standard choisis au hasard	Sujets 5 à 8 ans Niveau socioéconomique varié	MABC-2 Checklist	DCDQ	QUADAS-2
Asunta P et al, 2019 [37]	Etude Transversale Comparatif			DCDQ/MABC-2 notamment		
Ray-Kaeser S et al, 2019 [38]	Etude Transversale Comparatif	n=73 n=30 groupe clinique/n=43 groupe temoin	5 à 15 ans	DCDQ	MABC-2	QAREL QUADAS-2
Wilson B et al, 2009 [39]	Etude Transversale Comparatif	n=287	4 à 16 ans Niveau socioéconomique varié Pas d'handicap moteur et mental	DCDQ	MABC	QUADAS-2
Jaikaew R et al, 2019 [40]	Etude Transversale Comparatif	n=30	7 à 10 ans et 11 mois Aucune limitation physique et neurologique	MABC-2	BOT-2	QUADAS-2 QAREL
Ellinoudis T et al, 2011 [41]	Etude Transversale Comparatif	n=183 (98G/85F)	36 à 64 mois Pas d'handicap moteur ou mental	MABC-2		QAREL
Wuang Y et al, 2012 [42]	Etude Transversale Comparatif	n=144 (87G/57F)	6 ans à 6 ans 11 mois (n = 22) ;7 ans à 7 ans 11 mois (n = 22) ; 8 ans à 8 ans	MABC-2	PTPS	QAREL

			11 mois (n = 21) ; 9 ans à 9 ans 11 mois (n = 21) ;10 ans à 10 ans 11 mois (n = 20) ; 11 ans à 11 ans 11 mois (n = 19) ; et 12 ans à 12 ans 11 mois (n = 19) Pathologies associées et handicaps exclus			
Pannekoek L et al, 2012 [43]	Etude Transversale Comparatif	n=87 (35F/52G)	12 à 15 ans Pas d'handicap Statut socioéconomique varié	DCDQ	MABC-2	QUADAS-2
Tseng M et al, 2010 [44]	Etude Transversale Comparatif	n=1082	6 à 9 ans Principalement classe moyenne	DCDQ	MABC/BOT	QUADAS-2 QAREL
Rivard L et al, 2014 [45]	Etude Transversale Comparatif	n=3084	8 à 15 ans	DCDQ		
Cairney J et al, 2008 [46]	Etude Transversale Comparatif	n=523	9 à 14 ans	DCDQ	CSAPPA	QUADAS-2
Caravale B et al, 2015 [47]	Etude Transversale Comparatif	n=698	5 à 12 ans	DCDQ	MABC	QUADAS-2 QAREL

Caravale B et al, 2014 [48]	Etude Transversale Comparatif	n=78 Groupe clinique = 26	5 à 12 ans	DCDQ-IT	DCDQ	QUADAS-2 QAREL
Holm I et al, 2013 [49]	Etude Transversale Comparatif	n=45 (23F/22G) n=30 pour fiabilité inter-évaluateur n=29 pour fiabilité test-retest n=14 participants aux deux tests de fiabilité	7 à 9 ans et 11 mois	MABC-2		QAREL

3.2 Risques de biais des études incluses

3.2.1 Grille d'analyse utilisée

Deux grilles d'analyse ont été utilisées pour sélectionner les études à inclure dans la revue. Ce sont les échelles méthodologiques QUADAS-2 et QAREL. Elles définissent la qualité des études diagnostiques à inclure dans les revues systématiques en étudiant la validité pour l'une, la fiabilité pour l'autre. On synthétise les résultats à l'aide de tableaux.

3.2.2 Synthèse des biais retrouvés

Tableau 5 : Tableau présentant les résultats de l'échelle méthodologique QUADAS-2 pour les études diagnostiques de validité [24]

Etude	Risque de biais				Applicabilité		
Ferreira L, Gabbard C, Vieira J, da Silva P, Cheuczuk F, da Rocha F, de Souza V, Caçola P, Revista Brasileira de Medicina do Esporte, 2019 [32]	☺	☺	☹	☺	☺	☺	☺
Martini R, St-Pierre M, Wilson B, Canadian Journal of Occupational Therapy, 2011 [33]	☺	☹	☹	☺	☺	☺	☺
Montoro A, Capistrano R, Ferrari E, Da Silva Reis M, Cardoso F, Beltrame T, Journal of Human Growth and Development, 2016 [34]	☺	☺	☹	☺	☺	☺	☺
Capistrano R, Ferrari E, De Souza L, Beltrame T, Cardoso F, Motriz. Revista de Educacao Fisica, 2015 [35]	☺	☹	☹	☺	☺	☺	☺
Schoemaker M, Niemeijer A, Flapper B, Smits-Engelsman B, Developmental medicine and child neurology, 2012 [36]	☹	☺	☹	☺	☺	☺	☺
Ray-Kaeser S, Thommen E, Martini R, Jover	☹	☺	☹	☺	☺	☺	☺

M, Gurtner B, Bertrand A, PLoS ONE, 2019 [38]							
Wilson B, Crawford S, Green D, Roberts G, Aylott A, Kaplan B, Physical and Occupational Therapy in Pediatrics, 2009 [39]							
Jaikaew R, Satiansukpong N, Occupational Therapy International, 2019 [40]							
Pannekoek L, Rigoli D, Piek J, Barrett N, Piek J, Barrett N, Schoemaker M, Adapted Physical Activity Quarterly, 2012 [43]							
Tseng M, Fu C, Wilson B, Hu F, Research in Developmental Disabilities, 2010 [44]							
Cairney J, Missiuna C, Veldhuizen S, Wilson B, Human Movement Science, 2008 [46]							
Caravale B, Baldi S, Capone L, Presaghi F, Balottin U, Zoppello M, Research in Developmental Disabilities, 2015 [47]							
Caravale B, Baldi S, Gasparini C, Wilson B, European Journal of Paediatric Neurology, 2014 [48]							

Tableau 6 : Tableau présentant les résultats de l'échelle méthodologique QAREL pour les études diagnostiques de fiabilité [25]

Etudes	1	2	3	4	5	6	7	8	9	10	11
Martini R, St-Pierre M, Wilson B, Canadian Journal of Occupational Therapy, 2011 [33]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Ray-Kaeser S, Thommen E, Martini R, Jover M, Gurtner B, Bertrand A, PloS ONE, 2019 [38]	✓	✓	✓	✗	✗	✓	✓	?	✓	✓	✓
Jaikaew R, Satiansukpong N, Occupational Therapy International, 2019 [40]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Ellinoudis T, Evaggelinou C, Kourtessis T, Konstantinidou Z, Venetsanou F, Kambas A, Research in Developmental Disabilities, 2011 [41]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Wuang Y, Su J, Su C, Developmental Medicine and Child Neurology, 2012 [42]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Tseng M, Fu C, Wilson B, Hu F, Research in Developmental Disabilities, 2010 [44]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Caravale B, Baldi S, Capone L, Presaghi F, Balottin U, Zoppello M, Research in Developmental Disabilities, 2015 [47]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Caravale B, Baldi S, Gasparini C, Wilson B, European Journal of Paediatric Neurology, 2014 [48]	✓	✓	✓	✓	✓	✓	✓	?	✓	✓	✓
Holm I, Tveter A, Aulie V, Stuge B, Research in Developmental Disabilities, 2013 [49]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

3.3 Effets de l'intervention

3.3.1 Ferreira L et al, 2019

Résultats [32] :

- Le DCDQ présente une association au BOT-2 r=0,55 pour p<0,01
- On peut retirer de l'article les valeurs de sensibilité et de spécificité en fonction des scores de coupure : Se=75% et Sp=87% pour les 6 à 7 ans et 11 mois, Se=61,8% et Sp=93,8% pour les 8 à 9 ans et 11 mois, Se=54,5% et Sp=96,8 pour les 10 à 10 ans et 11 mois. On constate une très bonne spécificité et une sensibilité contrastée mais assez inférieure au seuil recommandé.
- Les scores de coupure sont en cours d'adaptation donc pas clairement définis. Le but de l'étude est de déterminer les scores de coupure les plus efficaces en comparant avec deux barèmes différents pour les comparateurs, un premier qui permet d'identifier avec une sécurité et une efficacité excellente (Se=100%) mais laisse beaucoup de faux négatifs et de patients non identifiés alors que le second permet d'identifier plus de sujets atteints mais avec moins de certitude.
- La sensibilité n'est pas assez importante (<75%), c'est-à-dire une mauvaise capacité à déterminer correctement les sujets atteints (les vrais positifs) et la spécificité est satisfaisante, c'est-à-dire une bonne capacité à déterminer correctement les sujets non atteints (les vrais négatifs)
- On peut voir des difficultés d'analyse car plus de brésiliens sont suspectés de TDC par rapport aux américains mais le niveau moteur brésilien est globalement moins bon que l'américain donc une recontextualisation est nécessaire

3.3.2 Martini R et al, 2011

Résultats [33] :

- On constate une mauvaise sensibilité : la sensibilité du DCDQ'07 (version en anglais) Se=42,8% et celle du DCDQ-FC Se=47,4%
- En revanche, la spécificité du DCDQ'07 (Sp=75,6%) et du DCDQ-FC (Sp=76,6%) est correcte
- Les deux questionnaires sont mieux à même d'identifier correctement les enfants qui n'ont pas de difficultés motrices que d'identifier les enfants atteints de TDC, c'est-à-dire qu'il est plus facile de discriminer justement les enfants non atteints des enfants atteints que l'inverse car la spécificité est bien meilleure que la sensibilité.
- Dans cette étude, seul le MABC-2 a été utilisé pour identifier les enfants comme étant TDC. Il faut être vigilant en ce qui concerne l'utilisation du MABC-2 comme seule mesure pour identifier les enfants atteints de TDC, bien qu'il soit parfois fait en milieu clinique.
- La valeur prédictive positive montre des résultats insuffisants (VPP=45% pour les deux versions)
- La valeur prédictive négative est correcte (VPN=73,9% pour le DCDQ'07 et VPN=78,3% pour le DCDQ-FC)

- L'élaboration de la traduction est rigoureuse et d'excellente qualité
- On observe une haute cohérence interne α=0,94 pour le DCDQ-FC et α=0,95 pour le DCDQ'07 traduisant l'homogénéité du questionnaire. Toutes les corrélations de chaque item avec le global étaient correctes (0,646 à 0,795) indiquant que tous les articles constituent des composantes de la coordination motrice. Le DCDQ'07 a démontré des caractéristiques similaires avec une bonne corrélation item-total (0.693 à 0.784). Plus précisément, les items 2 (attraper une petite balle), 12 (apprendre de nouvelles tâches motrices), 13 (ranger, attacher les chaussures, s'habiller) et 14 (ne jamais être décrit comme un « éléphant dans un magasin de porcelaine ») ont été ceux qui contribuent le plus à distinguer les TDC, tandis que les items 4 (saute facilement par-dessus les obstacles), 5 (court aussi rapidement que les autres) et 8 (écrire un dessin reconnaissable) sont les moins capables.
- Le coefficient ICC entre le MABC-2 et le DCDQ'07 était de 0.490 et le coefficient ICC utilisant le DCDQ-FC était 0.465, indiquant une validité concurrente passable et l'équivalence des versions française et anglaise mais le plus intéressant est que le DCDQ-FC a identifié 47% des mêmes enfants que le MABC-2 a identifié comme étant DCD et identifié 77% des mêmes enfants qui n'ont pas eu de problèmes mesurés par le MABC-2. Ces valeurs étaient similaires au DCDQ'07 lorsqu'il était utilisé avec ce groupe d'enfants.
- Il existe une excellente fidélité test-retest avec de bonnes corrélations intra classes : intra langue 0,916 et inter langue 0,990 mais aussi inter langue temporelle du français vers l'anglais 0,913 et de l'anglais vers le français 0,934
- Pour évaluer la fiabilité test-retest, un intervalle de 30 jours entre le test et le retest était visé, selon la disponibilité des parents pour la deuxième administration, cela a abouti à une fourchette réelle plus longue de 28 à 92 jours (moyenne = 39,38 jours) sans affecter la stabilité du questionnaire dans le temps par rapport à l'étude original de Wilson
- Le temps d'épreuve est rapide 10-15 minutes
- L'échantillon est peu représentatif car avec un effectif trop petit (n=84)

3.3.3 Montoro A et al, 2016

Résultats [34] :

- Dans l'évaluation du MABC-2, dans le groupe d'âge de 7 à 8 ans, 26,9% des filles et 34,4% des garçons étaient classés comme ayant un problème moteur alors que le questionnaire DCDQ –BR a donné des découvertes similaires. Pour le groupe d'âge de 9 à 10 ans, 21,4% des filles et 24,5% des garçons étaient classés comme ayant un problème moteur selon MABC-2 alors que 30,2% des filles et 22,3% des garçons étaient classés comme tels selon le DCDQ-BR
- Lors de la comparaison du MABC-2 et du DCDQ-BR pour les filles dans les groupes d'âge de 7-8 ans et 9-10 ans, on observe des corrélations positives entre le score total du MABC-2 et le DCDQ-BR selon le coefficient de Spearman ρ (7-8 ans : $\rho = 0,47$ et $p < 0,01$; 9-10 ans : $\rho = 0,32$ et $p < 0,01$). En ce qui concerne les garçons, les résultats montrent une corrélation positive significative entre le score total du MABC-2 et le

DCDQ-BR (7-8 ans : ρ = 0,62 et p < 0,01 ; 9-10 ans : ρ = 0,35 et p < 0,01). On peut ainsi constater une relation monotone croissante modérée entre 0,32 et 0,62.

- Une étude de validité simultanée est réalisée, en s'appuyant sur un étalon or le DCDQ selon les travaux de Wilson, avec pour objectif de montrer la capacité à établir le même diagnostic avec deux outils différents usités dans la pratique clinique quotidienne des praticiens. Cette étude contribue aux connaissances dans ce domaine en démontrant qu'il existe une corrélation positive et une relation entre les scores totaux MABC-2 et DCDQ-BR, indépendamment du sexe ou du groupe d'âge. Cela suggère une validité simultanée des deux instruments qui permet d'offrir une homogénéité de l'hypothèse diagnostic.
- En outre, en raison de l'association entre le DCDQ-BR et le MABC-2, ils peuvent servir comme un bon indicateur pour les enfants avec des troubles ensemble ou indépendamment, fournissant un diagnostic préliminaire qui peut être utilisé pour identifier des enfants qui ont besoin d'approches cliniques ou pédagogiques spécifiques.
- Le DCDQ est utilisé comme le gold standard dans cette étude alors que les preuves de ses qualités intrinsèques ne sont pas encore établies.

3.3.4 Capistrano R et al, 2015

Résultats [35] :

- L'échantillon n'est pas représentatif avec un petit effectif (n=40)
- Il est remarqué un manque de rigueur dans le processus de construction et de réalisation de l'essai (mauvais évaluateur : personne non qualifié utilise le MABC-2) donc pas de notion de validité à extraire de cette étude mais une donnée est intéressante à retirer concernant les différences entre les 3 catégories d'évaluateurs présentes : les enseignants, les enseignants d'éducation physique et les parents
- Dans l'évaluation faite par les enseignants grâce à la liste de contrôle MABC-2, 70,2% des sujets étaient qualifiés de « développement sain », 10,6% ont été classés à risque et 19,2% comme « problèmes moteurs définis ». Les enseignants d'éducation physique ont identifié 78,9% des enfants comme avec un « développement sain », 15,8% ont été classés comme à risque et seulement 5,3% ont été identifiés comme des « problèmes moteurs définis ». Pour les parents des enfants, 42,5% des enfants ont été identifiés comme ayant un « développement sain », 15,0% ont été classés comme à risque et 42,5% comme ayant des problèmes moteurs définis.
- Le test post hoc de Tukey a confirmé qu'il n'y avait pas de différence entre le jugement du comportement moteur des enfants dirigés par les parents et les enseignants (p = 0,24) mais en comparant les parents et les enseignants d'éducation physique, il y avait une différence statistiquement significative (p = 0,01) au même titre que quand on compare les enseignants et les enseignants d'éducation physique, il y avait également une différence statistiquement significative (p = 0,038). Ce test post-hoc (ou test de comparaisons multiples) peut être utilisé pour déterminer les différences significatives entre les moyennes de groupes dans une analyse de variance. On peut ainsi constater que les résultats des enseignants d'éducation physique sont significativement

différents de ceux des parents et des enseignants. Etant donné leurs compétences dans l'analyse du mouvement, la compréhension de l'acte moteur et leurs connaissances anatomiques ou théoriques, ces professionnels sont censés être plus qualifiés pour apprécier le comportement moteur de l'enfant avec fidélité et ainsi reconnaitre plus efficacement les troubles en se référant à la liste de contrôle du MABC-2 ici. Le kinésithérapeute étant un professionnel qualifié de l'acte moteur perturbé, on peut penser que ses observations seraient encore plus fines et pertinentes.

3.3.5 Schoemaker M et al, 2012

Résultats [36] :

- La méthode et la procédure sont bien expliquées et de très bonne qualité
- L'alpha de Cronbach était de 0,94 pour les 30 éléments ensemble, tous les éléments mesurent la même construction. La cohérence interne des éléments de la liste de contrôle était très bonne, ce qui implique que chaque item aborde un aspect de la performance motrice fonctionnelle.
- L'exploration de la structure factorielle de la liste de contrôle a révélé une solution à six facteurs, ce qui implique que la liste de contrôle mesure un large éventail de capacités motrices fonctionnelles. C'est un instrument conçu pour évaluer la déficience motrice fonctionnelle chez les enfants à risque de TDC compte tenu de l'hétérogénéité du trouble
- L'accord de classification entre la liste de contrôle et le test était de 80%
- Une corrélation modérée (ρ de Spearman) a été obtenue entre le TMS (Score Moteur Total) et le TTS (Score Total Test) (ρ=0,38 ; p <0,001 ; n = 383) et entre les scores totaux du DCDQ'07 et du TMS (ρ = 0,36 ; p <0,001 ; n = 130). Le pourcentage d'accord entre la liste de contrôle et le test était de 80% (307⁄383). Une valeur kappa de 0,28 était obtenue, ce qui est considéré comme une légère concordance.
- La sensibilité était de 41% et la spécificité 88% dans tous les groupes d'âge si le 15ème centile a été utilisé comme seuil pour le MABC-2. Un taux bas pour la sensibilité ou la spécificité est une constatation courante dans les échantillons de population.
  Le MABC-2 fournit des données quantitatives objectives relatives aux compétences psychomotrices. Une note standard pour chacun des trois domaines est fournie ainsi qu'une note totale de test, reflétant la performance motrice globale de l'enfant.
  Un système d'alerte basé sur des seuils critiques (percentiles correspondants) indique si la note totale de test obtenue le situe en zone rouge « trouble moteur manifeste » (5ème centile), en zone orange « à risque » (15ème centile) ou zone verte « performance dans la norme ». L'examen des résultats aux différents domaines permet d'établir un profil plus précis des difficultés de l'enfant. Les données qualitatives recueillies au cours de l'examen viennent enrichir cette analyse.
- ASC = 0,673. Plus l'Aire Sous la Courbe ROC est proche de 1, plus le test est performant.
- Dans l'ensemble, 16,7% (n = 64) des enfants ont été évalués en dessous du 15e centile sur les deux instruments (Checklist et Test), et 6,8% au-dessous du 5e centile pour le test et 5,7% pour la liste de contrôle.

- La liste de contrôle du MABC-2 est administrée aux enfants par les enseignants qui ne sont pas formés pour ça, il peut donc y avoir une probable erreur d'administration

3.3.6 Asunta P et al, 2019

Résultats [37] :

- C'est une revue systématique s'intéressant à la quasi-totalité des outils diagnostics des TDC dont le DCDQ et le MABC-2.
- La validité prédictive est meilleure sur une population à risque que sur la population globale.
- Le DCDQ, développé au Canada, avait le plus d'adaptations culturelles dans différents pays et il a relativement plus de tests psychométriques que les autres outils inclus dans la revue, c'est ainsi le questionnaire le plus étudié et évalué. Un outil clinique valide, mais pas pour la population globale pour un dépistage mais plutôt pour un diagnostic dans une population cible de sujets suspectés. On ne relève aucun résultat de fiabilité inter-évaluateur et peu de conclusions satisfaisantes sur la validité apparente.
- Les forces du MABC-2 sont qu'ils existent quelques bons scores de fiabilité test – retest, il est traduit dans de nombreux pays. Ils restent des éléments à développer, on observe une très faible sensibilité (aucune des études ne répondait aux exigences). Des critères sont encore à évaluer comme la fiabilité inter-évaluateur non étudiée.
- Dans le futur, la propriété psychométrique des tests devrait être améliorée en tenant compte des descriptions de fonctionnement et d'utilisation. De plus, la stabilité (fiabilité test-retest) et l'homogénéité (fiabilité inter évaluateur) devraient être évaluées davantage lors de l'évaluation des propriétés psychométriques d'un instrument diagnostic.

3.3.7 Ray-Kaeser S et al, 2019

Résultats [38] :

- Les scores de coupure sont à définir au début de l'étude en raison de la nécessaire adaptation interculturelle
- Le petit nombre de participants est un défaut pouvant rendre difficile l'interprétation des statistiques relevant de la validité prédictive
- La cohérence interne est excellente avec l'alpha de Cronbach $\alpha = 0,96$. Les coefficients alpha de Cronbach, lorsque chaque élément a été supprimé, indiquaient qu'aucun élément posait un problème ($\alpha = 0,95–0,96$) et ainsi que chaque item mesure la même construction.
- Tous les éléments étaient moyennement à fortement corrélés avec le score total ($r = 0,673$ à $0,836$)
- La fiabilité du retest est bonne après 38 jours de moyenne entre les deux épreuves avec un écart type de 17 jours pour un objectif de 30 jours. En effet, le coefficient intra classe (avec un intervalle de confiance à 95%) était 0,865 pour le groupe clinique ($n = 28$), 0,814 pour le groupe témoin ($n = 42$) et 0,956 pour l'échantillon total ($n = 70$), sans différence entre les scores aux deux points de mesure différents ($p > 0,05$), indiquant une bonne fiabilité.

- La corrélation obtenue entre les scores DCDQ-FE et MABC-2 d'après le coefficient de corrélation ρ de Spearman ρ = 0,802 avec p < 0,001 prend en charge la validité convergente
- La sensibilité est de 85% et la spécificité est de 81,6% avec une aire sous la courbe de 0,896 indiquant une bonne capacité pour détecter les enfants à risque de TDC et un test performant en plus des critères de fiabilité et validité précédemment décrits
- Les valeurs prédictives positives et négatives étaient respectivement de 0,71 et 0,91. La valeur prédictive positive (VPP) est la probabilité que le patient, dont le test est positif, soit effectivement malade. La valeur prédictive positive correspond donc au nombre de personnes malades et positives au test (vrais positifs) parmi l'ensemble des personnes positives au test. La valeur prédictive négative (VPN) est la probabilité que le patient, dont le test est négatif, ne soit pas malade. La valeur prédictive négative correspond donc au nombre de personnes non-malades et négatives au test (vrais négatifs) parmi l'ensemble des personnes négatives au test. Contrairement à la sensibilité et à la spécificité, les valeurs prédictives du test ne sont pas constantes, mais dépendent de la prévalence de la maladie dans la population. Les valeurs de prédiction mesurent donc les pourcentages de personnes correctement classées en fonction du résultat donné par un test de dépistage. Dans une optique de dépistage, la VPP donne une idée du nombre d'examens inutiles qui seront faits à une population en bonne santé.

3.3.8 Wilson B et al, 2009

Résultats [39] :

- Le coefficient α de Cronbach pour le questionnaire révisé était de 0,94, et le coefficient α, si un des 15 éléments était supprimé, était supérieur à 0,93 indiquant tous les éléments mesurent la même construction. Les corrélations corrigées entre les éléments individuels et le score total variaient de 0,52 à 0,78 indiquant une forte cohérence interne.
- Les résultats de la régression logistique ont indiqué que le statut TDC était prédit de manière significative par le score DCDQ révisé (p < 0,001)
- L'étude souhaite objectiver les scores de coupure les plus performants pour identifier le plus de patients parmi les participants et après les avoir trouvés pour chaque tranche d'âge, il est possible d'observer les statistiques relatant de la validité prédictive. Pour les enfants de moins de 8 ans, on obtient une sensibilité de 75% et spécificité de 71%. Pour les enfants entre 8 et 10 ans, la sensibilité était 89% et la spécificité était de 67%. Pour les enfants de plus de 10 ans, la sensibilité est de 89% et la spécificité de 76%. Pour l'échantillon global, si les seuils spécifiques à l'âge sont utilisés, la sensibilité et la spécificité sont augmentées à 85% et 71% respectivement. On peut considérer la sensibilité comme satisfaisante, le seuil étant souvent défini à 80% notamment par l'APA, alors que la spécificité est assez inférieure aux recommandations statistiques de 90%. L'effectif étant plutôt petit, il est dur d'extrapoler ces valeurs à la population globale d'autant plus qu'une grande partie de l'échantillon était précédemment

suspectée de TDC or on sait que la validité prédictive est meilleure dans une population cible que dans la population générale.

- Trois groupes d'âge ont émergé dans le descriptif analytique des scores à tous les âges, et ce groupe a entraîné une sensibilité accrue lorsque les courbes ROC ont été examinées : les groupes étaient composés d'enfants de moins de 8 ans, de 8 ans à 9 ans 11 mois et 10 à 15 ans. Il n'y avait aucune association significative entre le groupe d'âge et le sexe ainsi les filles et les garçons ne différaient pas sur leur score DCDQ révisé total. Cependant, l'âge était corrélé au score total ($r = 0,242$, $p < 0,001$). Ces résultats indiquent que l'échelle est valable pour une utilisation avec les deux sexes mais que des scores de coupure spécifiques à l'âge sont nécessaires.
- Pour ce qui concerne la validité concurrente avec le MABC première édition, les scores totaux pour le DCDQ révisé étaient corrélés avec les scores de déficience totale du MABC ($r = - 0.55$; $p < .001$). La corrélation avec le MABC est négative de façon appropriée car les deux tests sont échelonnés dans des directions opposées, en effet, les scores élevés de déficience MABC reflètent une mauvaise performance.
- Cette étude de Wilson est la base de toutes les études suivantes concernant ce DCDQ nouvelle formule et révisé pour plus de facilité d'administration et une meilleure efficacité.

3.3.9 Jaikaew R et al, 2019

Résultats [40] :

- La fiabilité inter-évaluateurs du MABC-2 variait de 0,71 à 1,00

3.3.10Ellinoudis T et al, 2011

Résultats [41] :

- Le coefficient intra classe est entre 0,61 à 0,96 pour chaque élément entre les deux épreuves avec un intervalle de confiance de 95%.
- La cohérence interne de la batterie a été testée en utilisant l'alpha de Cronbach. Les valeurs de dextérité manuelle, de visée et les catégories de capture et d'équilibre étaient respectivement de 0,51 ; 0,66 et 0,70, ce qui indique une cohérence interne plutôt modérée.
- Les corrélations étaient modérément significatives entre chacun des éléments du test et le score total, allant de $r = 0,35$ ($p < 0,05$) à $r = 0,62$ ($p < 0,05$). Dans l'étude actuelle, les trois composantes motrices étaient significativement corrélées entre elles mais seulement à un degré modéré mais ce résultat était attendu car les trois sous-échelles sont censées mesurer différents aspects du mouvement. Si une corrélation trop élevée existe, les sous-échelles mesureraient les mêmes caractéristiques. Inversement, si les corrélations étaient faibles, les sous-ensembles mesureraient des compétences non pertinentes. Les résultats de la présente étude concordent notamment avec les conclusions de Henderson et al, 2007, et Simons et al, 2008 qui ont signalé que des coefficients modérés sont acceptables parce que le MABC-2 a été conçu pour couvrir un large éventail de capacités motrices. Le fait que les composantes soient bien corrélées avec le score total fournit plus de preuves pour la validité du test.

3.3.11 Wuang Y et al, 2012

Résultats [42] :

- La cohérence interne du score total du test MABC-2 est excellente avec un α de Cronbach = 0,90. Les valeurs de Cronbach pour les sous-échelles de dextérité manuelle, de visée et capture et d'équilibre étaient respectivement de 0,81, 0,84 et 0,88. Cela implique une homogénéité suffisante de tous les domaines individuels ainsi que le test total.
- Dans l'analyse test-retest, les coefficients intra classe varient entre 0,88 et 0,99 pour les articles et les sous-échelles, alors que l'ICC du score total était de 0,97, ce qui indique une très bonne aptitude.

3.3.12 Pannekoek L et al, 2012

Résultats [43] :

- Un alpha de Cronbach élevé a été trouvé pour le DCDQ'07 (α =0,95). Le coefficient alpha n'augmentait pas de manière significative si l'un des éléments était supprimé, indiquant qu'aucun des éléments ne semblait poser de problème. De plus la suppression d'un élément ne consoliderait pas le DCDQ'07. La cohérence interne des sous-échelles a également été jugée élevée : α = 0,94 ; α = 0,88 ; α = 0,85.
- Les corrélations totales des éléments corrigés étaient toutes significatives et positives ($p < 0,001$), de r = 0,62 à r = 0,82. Des corrélations positives passables à modérées ont été trouvées entre les éléments, r = 0,31 à r = 0,86, tous atteignant une signification ($p < 0,001$). Les corrélations entre les sous-échelles étaient bonnes sans être trop élevées démontrant la cohérence des mesures et la diversité des aptitudes motrices. Ainsi, la corrélation entre le contrôle pendant le mouvement et la motricité fine/écriture manuscrite est de r =0,52, entre le contrôle pendant le mouvement et la coordination générale est de r = 0,77 alors que celle entre la coordination générale et la motricité fine/écriture manuscrite est de r = 0,73. Les corrélations entre les éléments appartenant à la même sous-échelle étaient en moyenne supérieures aux corrélations entre les éléments de différentes sous-échelles. Les articles appartenant à la même sous-échelle du DCDQ'07 étaient plus étroitement liés que les éléments de différentes sous-échelles. Néanmoins, les corrélations entre tous les éléments individuels étaient importantes. Cela suggère que les éléments mesurent une construction similaire.
- Une corrélation moyenne mais statistiquement significative a été trouvée entre les scores totaux sur le MABC-2 et le DCDQ'07 ($\rho = 0,34$; $p < 0,001$)
- Une aire sous la courbe de 0,81 a été obtenue (IC à 95% qui dépasse 100% et n'est donc pas réaliste) avec une sensibilité élevée du DCDQ'07 de 85,7% et une spécificité de 77,5%. La valeur prédictive positive était de 25% en conséquence du grand nombre de faux positifs dans l'identification des différences motrices difficultés selon le DCDQ'07. Une valeur prédictive négative de 98,4% a été atteinte. Il n'y a pas d'accord complet dans l'identification des difficultés motrices dans cet échantillon d'adolescents à l'aide du DCDQ'07 et du MABC-2. En effet le DCDQ'07 n'a pas identifié un des adolescents qui avait été identifié avec des difficultés motrices du MABC-2. Sur

les 24 participants identifiés avec des difficultés motrices par le DCDQ'07, seuls six ont été identifiés par le MABC-2. Une explication plausible des incohérences dans les résultats du test pourrait être que les mesures diffèrent par leur sensibilité autour du point de coupure.

3.3.13 Tseng M et al, 2012

Résultats [44] :

- L'analyse de la cohérence interne a été réalisée initialement sur le questionnaire avec 17 items comme conçu en utilisant l'échantillon entier (n = 1082). Le coefficient α de Cronbach a été calculé (α = 0,84), indiquant une cohérence interne élevée et l'homogénéité des éléments. Ensuite, une corrélation corrigée article-total a été calculée pour examiner la relation entre chaque élément et l'ensemble du questionnaire. Tous les coefficients étaient élevés à l'exception du point 14 (r = 0,01) et du point 11, qui était corrélée négativement (r = -0,35). Par conséquent, des articles 11 et 14 ont été supprimés, faisant passer le α à 0,89. Dans toutes les analyses ultérieures, la version à 15 éléments a été utilisée.
- La corrélation de Pearson a été calculée pour évaluer la fiabilité test-retest du DCDQ-C, avec un échantillon de 35 parents. Le coefficient de Pearson était de 0,94 (p <0,001), ce qui indique une grande fiabilité.
- Lorsque le MABC combiné avec le BOTMP a servi de critère standard et un point de coupure au 10ème centile pour le DCDQ-C a été appliqué donnant une sensibilité de 73% et une spécificité de 54%.

3.3.14 Rivard L et al, 2014

Résultats [45] :

- Pour l'échantillon de population (très grand échantillon), on observe des différences statistiques entre les sexes. En effet, les scores moyens au DCDQ pour les filles étaient supérieurs à ceux des garçons. Pour le groupe TDC, des différences de sexe significatives dans la moyenne des scores moyens au DCDQ ont été trouvées. Des scores moyens plus élevés ont été trouvés pour les filles mais de façon plus marquée et beaucoup plus grande que celle trouvée dans l'échantillon de la population.
- Il existe également des différences de moyenne au DCDQ selon l'âge qui étaient significatives, les enfants plus âgés démontrant des capacités plus élevées que les plus jeunes. Un test post hoc de Tukey a indiqué que la moyenne du groupe plus jeune (moins de 10 ans) était significativement différente que les moyennes pour les enfants de 10 à 11 ans et 11 mois (p < 0,05) et le groupe supérieur ou égal à 12 ans (p < 0,01). Il n'y avait pas de différence statistiquement significative entre les moyennes des deux groupes plus anciens (p > 0,5). L'analyse des scores par tranche d'âge a révélé des résultats non significatifs. Aucune sous-analyse supplémentaire par âge et sexe n'a été effectuée dans le groupe TDC en raison de la petite taille de l'échantillon.

3.3.15 Cairney J et al, 2008

Résultats [46] :

- L'alpha de Cronbach pour le DCDQ est élevé (α = 0,94). Chacune des sous-échelles a montré une cohérence interne également élevée : contrôle pendant le mouvement (α = 0,91), moteur fin / écriture manuscrite (α = 0,91) et coordination générale (α = 0,84). Les corrélations entre les sous-échelles étaient modérées à élevées (contrôle du mouvement et motricité fine, r = 0.58 ; contrôle du mouvement et coordination générale, r = 0,75 ; motricité fine et coordination générale, r = 0,68)
- La corrélation entre les deux échelles totales, le DCDQ et le CSAPPA, est modérée (r = 0,38, p <0,001). Les corrélations entre le score total du DCDQ et ceux de ses sous échelles sont excellentes : contrôle pendant le mouvement r = 0,90 ; motricité fine r = 0,84 ; coordination générale r = 0,92

3.3.16Caravale B et al, 2015

Résultats [47] :

- Le coefficient α de Cronbach pour le score total DCDQ est bon 0,89 ainsi que ceux des sous-échelles : le contrôle pendant le mouvement (0,83), la motricité fine et l'écriture manuscrite (0,79) et la coordination générale (0,73).
- La stabilité du DCDQ a été considérée sur un plus petit échantillon de participants (n = 45) pour le test-retest. Le coefficient intra classe est excellent pour le score total du DCDQ (0,88 ; p < 0,01) au même titre que pour le contrôle pendant le mouvement (0,87 ; p < 0,01), tandis que pour les deux autres facteurs les résultats étaient moins bons la motricité fine/écriture manuscrite (0,48 ; p < 0,01) et la coordination générale (0,40 ; p < 0,01).
- Pour l'échantillon total avec un score de coupure de 53 ou moins, soit le même score rapporté par échantillon au 15e centile, avec une sensibilité de 59% et une spécificité de 65%. L'aire sous la courbe est de 0,62. Des différences par rapport au DCDQ'07 sont devenus évidents lorsque l'âge des enfants a été pris en compte : dans le groupe des 5 à 7 ans, le seuil pour l'échantillon italien était de 52 ou moins au 15e centile avec une sensibilité de 67% et une spécificité de 54% (ASC = 0,59). Pour le groupe des 8 à 10 ans, le point de coupure pour l'échantillon italien était de 52 ou moins au 15e centile avec une sensibilité de 73% et une spécificité de 70% (ASC = 0,71) ; enfin pour le groupe des 11 à 12 ans, le point de coupure pour l'échantillon italien, nous avons trouvé un point de coupure de 56 ou moins au 15e centile avec une sensibilité de 58% et une spécificité de 67% (ASC = 0,59). Caravale et son équipe cherche à définir les scores de coupures les plus efficaces pour la population italienne et obtenir la meilleure validité prédictive mais elle demeure assez médiocre et inférieure à ses homologues étrangers ou l'étude initiale de Wilson.

3.3.17Caravale B et al, 2014

Résultats [48] :

- La cohérence interne a été calculée en considérant l'ensemble de l'échantillon de 78 enfants. L'α de Cronbach était de 0,94 pour le score total, avec un coefficient α allant de 0,93 à 0,94 si les éléments ont été supprimés. La corrélation item-article total variait de 0,58 à 0,82.

- La fiabilité test-retest n'a été calculée que sur un échantillon de 24. La fiabilité des tests pour chaque élément a montré que 11 éléments avaient des coefficients intra classes élevés allant de 0,82 à 1, trois éléments avaient les indices modérés (0,62 à 0,75) et que seulement un avait un ICC faible (0,43)
- On calcule une sensibilité de 88% et une spécificité de 96% indiquant une bonne capacité de dépistage (mais ces valeurs sont trouvées dans des populations cibles de positifs pour la sensibilité et de négatifs pour la spécificité) ainsi qu'une valeur prédictive positive de 82% et une valeur prédictive négative de 90%.

3.3.18 Holm I et al, 2013

Résultats [49] :

- La fiabilité inter-évaluateurs est jugée sur la même journée avec deux évaluateurs indépendants et en aveugle. Le coefficient de corrélation inter classe est de 0,35 à 0,67 avec un intervalle de confiance à 95%.
- La fiabilité test-retest est évaluée par le même évaluateur avec une à deux semaines entre les deux épreuves. Le coefficient de corrélation intra classe est de 0,23 à 0,76 avec un intervalle de confiance à 95%. Il y a deux items avec des résultats pas assez importants : le saut sur un tapis avec la jambe droite ou avec la jambe gauche 0,24 ; attacher ses lacets 0,35. Les autres items sont supérieurs à 0,48.

4 Discussion

Cette revue avait pour intention de déterminer au travers de la littérature scientifique, sur les dix dernières années, la qualité des instruments diagnostics que sont le DCDQ et le MABC-2, c'est-à-dire leur validité et leur fiabilité. On s'intéresse ainsi aux études travaillant sur le sujet avec pour objectif de définir s'il est possible de s'en servir avec efficacité et sécurité dans le cadre d'une suspicion de troubles moteurs pouvant se rapporter aux TDC et notamment par le kinésithérapeute au cours de la thérapie de l'enfant. Cela permet alors de confirmer la valeur scientifique des outils par les preuves pour son utilisation clinique et d'offrir au kinésithérapeute la possibilité de participer à l'opération diagnostic et ainsi de réduire le nombre d'enfants non diagnostiqués. En effet, la littérature s'accorde sur l'importance capitale d'une identification précoce mais est rendue délicate par l'absence d'un critère standard accepté et démontré pour les TDC et d'un test diagnostic scientifiquement établi et justifié. L'évaluation multiple prend ainsi tout son sens et il est alors pertinent d'étudier deux instruments différents et complémentaires comme le sont le DCDQ et le MABC-2.

Ainsi, l'interprétation des principaux résultats s'est faite sur des paramètres statistiques relatifs à la validité et à la fiabilité d'une étude diagnostique.

4.1 Analyse des principaux résultats

On remarque que les articles précédemment présentés abordent de nombreux thèmes en relation avec les TDC, le MABC-2, le DCDQ et les études diagnostiques. On peut alors constater plusieurs informations importantes en s'appuyant sur les résultats observés.

On peut s'intéresser à la population étudiée. C'est une population d'enfants de 5 à 15 ans pour le DCDQ et de 3 à 16 ans pour le MABC-2. La majorité du temps, les adolescents ne sont pas inclus dans l'étude ([32],[34],[35],[36],[40],[41],[42],[44],[47],[48]), les recherches se concentrant plutôt sur un groupe de sujets plus jeunes de 5 à 12 ans, patients le plus souvent vus dans la situation clinique pour un diagnostic. On peut voir que les effectifs comportent d'autres différences. Les tailles d'échantillons sont variées et parfois trop petites ([33],[35],[38],[40],[43],[48]) rendant compliquées les conclusions et leur transposition à la population générale. En effet, il est complexe de réaliser des études de grande échelle, de trouver un échantillon de population représentatif et de taille suffisante.

De plus, il existe de nombreuses adaptations interculturelles pour pouvoir exporter et démontrer les outils diagnostics à un niveau mondial. Ainsi, la littérature s'efforce à recalibrer les instruments à la population locale du pays en retravaillant les scores de coupure pour le DCDQ par exemple, mais aussi en adaptant les items de chacun des outils à la culture du pays et/ou aux expressions caractéristiques d'un pays, par exemple pour exprimer la maladresse en Chine, il est évoqué un taureau dans un commerce alors qu'en France on parlerait d'un éléphant dans un magasin de porcelaine. On constate également que le questionnaire diffère entre les versions francophones canadienne et européenne montrant les différences culturelles ou bien encore qu'un acte moteur présent et commun dans un pays ne l'est pas dans l'autre, en cherchant à afficher la meilleure traduction possible pour la plus grande compréhension possible. Le niveau moteur global de la population d'un pays n'étant pas le

même d'un à l'autre, le travail d'ajustement est nécessaire, le Brésil a par exemple un niveau moteur inférieur à celui des Etats-Unis, il est donc difficile de faire une comparaison directe entre les deux résultats. L'élaboration rigoureuse de la traduction est de ce fait un critère de qualité important de l'étude, celle de Martini et al, 2011 en est un exemple pour les francophones du Canada au même titre que celle de Jaikaew et al, 2019 pour les enfants thaïlandais parmi d'autres.

On peut également s'apercevoir qu'à l'intérieur de ces échantillons, il existe des différences de résultats entre les sexes et entre les âges. On peut alors observer les différentes compétences évaluées par le MABC-2, où la motricité fine et l'écriture manuscrite va être une aptitude plutôt féminine avec des meilleurs résultats pour elles, alors que les lancers, la course ou encore les sauts vont plutôt être du domaine des garçons. Il s'agit dès lors de considérer les différences entre les sexes et s'interroger si intégrer des normes pour les distinguer permettrait une meilleure discrimination entre sujets atteints et non atteints. On peut voir une sur-identification des garçons et une sous-identification des filles. En fonction du sexe, les résultats seront différents également en ce qui concerne le DCDQ avec généralement une meilleure évaluation des filles que celle des garçons. Ainsi les scores moyens au DCDQ sont supérieurs pour les filles par rapport à ceux des garçons selon Rivard et al, 2014 mais c'est à mettre en relation avec la prévalence des garçons dans la population présentant des TDC supérieure à celle des filles selon l'Inserm dans sa synthèse de l'expertise collective. Wilson et al, 2009, étude originelle et base de toute la recherche sur le DCDQ révisé, a permis de montrer, en cherchant à objectiver si des scores de coupure offraient une meilleure efficacité et lesquels seraient alors ces derniers, qu'il existait aucune association significative entre le groupe d'âge et le sexe, ainsi les filles et les garçons ne différaient pas sur leur score total au DCDQ révisé. Cependant, l'âge était corrélé au score total ($r = 0,242$; $p < 0,001$). Ces résultats indiquent que l'échelle est valable pour une utilisation avec les deux sexes mais que des scores de coupure spécifiques à l'âge sont nécessaires. Wilson les définit en utilisant les performances de la validité prédictive, c'est-à-dire la sensibilité et la spécificité, et les délimite de 5 ans à 7 ans et 11 mois puis de 8 ans à 9 ans et 11 mois et enfin de 10 à 15 ans. Wilson et son équipe en sont arrivés à cette conclusion concernant les scores de coupure en s'appuyant sur les différentes valeurs de sensibilité et de spécificité en fonction de la détermination des seuils. En effet, en utilisant les scores de coupure, on obtient pour les enfants de moins de 8 ans (Se = 75% ; Sp = 71%), pour les enfants de 8 à 10 ans (Se = 89% ; Sp = 67%), pour les enfants de plus de 10 ans (Se = 89% ; Sp = 76%) ce qui donne une sensibilité de 85% et une spécificité de 71% pour l'échantillon global alors que sans se servir des seuils, on trouve une sensibilité de 81% et une spécificité de 65%. On constate ainsi une meilleure capacité statistique de détermination des sujets atteints ou non.

On note par ailleurs des différences statistiques concernant la validité prédictive en fonction de la population étudiée, soit un échantillon de la population globale, soit un échantillon clinique. Il faut comprendre que le diagnostic se distingue du dépistage par une caractéristique fondamentale : la motivation de l'examen. S'il est réalisé en raison de l'état clinique (sujet suspect), c'est un diagnostic, alors que s'il est réalisé indépendamment de l'état clinique (sujet apparemment sain), c'est un dépistage. Ici, les études portent principalement sur sa capacité à déterminer les sujets sains et les sujets malades dans une population globale alors que le

but du kinésithérapeute est de porter un premier avis éclairé et scientifiquement prouvé sur un patient suspect quant à son développement moteur. Or, on sait que les résultats et l'efficacité sont meilleurs quand le test s'adresse à une population à risque plutôt qu'à la population générale. On remarque que la sensibilité et la spécificité dans l'étude de Caravale et al, 2014 sont excellentes (Se = 88% et Sp = 96%) à l'instar de l'étude de Ray-Kaeser et al, 2019 (Se = 85% ; Sp = 81,6 ; ASC = 0,896) alors qu'elles possèdent toutes les deux un groupe témoin et un groupe clinique augmentant ainsi les potentiels de la validité prédictive. La sensibilité ici est bonne et répond aux critères seuils de l'APA d'au moins 80% dans les deux cas alors que la spécificité est bonne pour l'étude de Caravale et légèrement insuffisante pour celle de Ray-Kaeser car le seuil est de 90% pour la spécificité selon l'APA. Ces dernières s'intéressent au DCDQ.

La sensibilité fait référence au pourcentage d'enfants avec des troubles moteurs qui est correctement détecté (80% est un seuil acceptable selon l'APA), c'est l'aptitude d'un test à identifier correctement les individus malades grâce à une réponse positive. La spécificité se réfère au pourcentage d'enfants sans problèmes correctement identifié (90% est préférable), c'est l'aptitude d'un test à identifier correctement les individus non malades grâce à une réponse négative. La validité prédictive peut ainsi être mesurée et évaluée. La sensibilité du DCDQ est globalement insuffisante dans la population globale comparée à la norme alors que la spécificité est généralement satisfaisante dans cette même population. La sensibilité varie entre 42,8 et 85,7% alors que la spécificité quant à elle oscille entre 54 et 96,8%. On peut remarquer que la sensibilité et la spécificité sont insatisfaisantes et inférieures aux valeurs obtenues dans un groupe clinique mais avec une spécificité majoritairement meilleure que la sensibilité. Cette hétérogénéité dans les résultats obtenus peut être expliquée par les différences de taille d'échantillon. En effet, il est dur d'extrapoler la validité prédictive à une population générale quand elle est évaluée sur un échantillon petit et alors trop dépendant des sujets intégrés dans l'étude. Les études avec un effectif trop court sont ainsi trop sensibles à de multiples variations. Le manque d'exactitude et d'uniformité dans la définition de la validité prédictive des outils diagnostics des TDC est une limite de la validation de ceux-ci et est responsable de l'incertitude en moyen de diagnostic unique du DCDQ et du MABC-2, ce qui implique un diagnostic pluriel et ainsi plus coûteux mais plus efficace. L'aire sous la courbe des études se situe entre 0,59 et 0,896. Un seuil est idéal s'il permet de séparer totalement les positifs des négatifs, sans faux positifs ni faux négatifs, c'est-à-dire ASC = 1. Ainsi, plus l'ASC est proche de 1 plus le test est performant. Les résultats relevés ici démontrent une efficacité modérée et sans doute insuffisante pour un diagnostic efficace par son utilisation seule. L'étude de Pannekoek et al, 2012 est la plus efficace statistiquement concernant la validité prédictive. Les valeurs prédictives positives et négatives sont aussi intéressantes à relever. Ces valeurs ne sont pas constantes et dépendent de la prévalence. La VPP donne une idée du nombre d'examens inutiles qui seront faits à une population en bonne santé. La VPP varie entre 25 et 82% et la VPN entre 73,9 et 98,4%. La seule étude sur la validité prédictive du MABC-2 est celle de Schoemaker et al, 2012. Elle démontre une mauvaise sensibilité (41%) et une bonne spécificité (88%) alors que l'aire sous la courbe est de 0,673. Cependant, le manque de recherche sur ce sujet ne permet pas d'avoir un avis définitif, il faut continuer à étudier la validité prédictive du MABC-2 mais pour l'instant à l'instar du DCDQ, il ne semble pas pouvoir

diagnostiquer seul les TDC. Il s'agit d'être vigilant quant à l'utilisation seule de l'un ou l'autre même si cela est souvent fait en clinique par les professionnels pour le MABC-2 qui, de son côté, s'attarde sur les capacités motrices directement. Le DCDQ est un questionnaire destiné aux parents qui ne s'attachent pas à traiter les mêmes choses, s'attachant à observer les empêchements moteurs du quotidien. Ils sont donc assez complémentaires pour une évaluation ensemble. Les critères évalués par ces instruments ne sont, par ailleurs, pas les mêmes. Le DCDQ traite le critère B du DSM-V au même titre que la liste de contrôle du MABC-2 alors que le test moteur du MABC-2 s'occupe du critère A selon Schoemaker et al, 2012. Le critère A concerne les difficultés dans les habiletés motrices coordonnées alors que le critère B concerne la présence persistante dans les activités de la vie courante et les désagréments que cela engendre. Les outils diagnostics sont corrélés ou associés entre eux mais également avec d'autres tests d'évaluation diagnostic. Il existe ainsi une cohérence positive et une relation entre les scores totaux MABC-2 et le DCDQ-BR selon Montoro et al, 2016, indépendamment du sexe et du groupe d'âge (entre r = 0,32 et 0,62). Cela suggère une validité simultanée des deux outils et offre une homogénéité de l'hypothèse diagnostique ainsi qu'une association pertinente. Dans l'étude de Martini et al, 2011, le DCDQ'07 est correctement associé au MABC-2 (r = 0,49) et le DCDQ-FC est associé avec la même puissance au MABC-2 (r = 0,465). Le travail de Pannekoek et al, 2012, montre une corrélation correcte (r = 0,34 ; p < 0,001) entre le MABC-2 et le DCDQ'07. Ray-Kaeser et al, 2019, démontre une excellente association entre le DCDQ et le MABC-2 (r = 0,802 ; p < 0,001). Le DCDQ est également corrélé avec deux autres tests appelés BOT-2 et MABC, positivement pour le BOT-2 (r = 0,55 ; p < 0,01) d'après l'étude de Ferreira et al, 2019 et négativement pour le MABC, première version de la batterie d'évaluation du mouvement de l'enfant, selon Wilson et al, 2009, (r = - 0,55 ; p < 0,001). Le MABC ayant un système de notation opposé, la relation négative entre lui et le DCDQ est normale. Le DCDQ est également modérément corrélé avec le CSAPPA (r = 0,38 ; p < 0,001) selon Cairney et al, 2008. Enfin le dernier article de Schoemaker et al, 2012 évoque l'association entre les deux parties du MABC-2, la liste de contrôle et le test sont corrélés positivement (r = 0,38 ; p < 0,01) alors que la liste de contrôle et le DCDQ'07 le sont également avec la même puissance (r = 0,36 ; p < 0,01).

On rencontre plusieurs problèmes dans ces associations entre tests diagnostics. Premièrement, l'élément étudié, que ce soit le DCDQ et le MABC-2, est majoritairement comparé avec un gold standard non démontré. En effet, les tests pour les TDC ne sont pas scientifiquement établis pour être considérés comme un étalon d'or dans la détection des troubles moteurs. De plus, le premier est comparé à de multiples reprises au second ou inversement. Il est ainsi délicat d'affirmer le potentiel de la qualité diagnostique d'un test en le comparant à un autre instrument qui est justement prouvé dans d'autres études en le comparant à ce premier test. Ainsi aucun ne semble être démontré et justifié par des preuves solides. Dans l'étude de Wilson en 2009, le questionnaire DCDQ est révisé et comparé au MABC, un des seuls tests à être validé mais trop long, coûteux et compliqué à mettre en place justifiant la création de la deuxième version, en montrant une bonne association entre les deux et une validité prédictive plutôt bonne. Il est aussi mis en comparaison avec le BOT-2, lui aussi valide, avec une bonne corrélation également. Le MABC-2 découlant du MABC, test valide, il est souvent dit que la validité de ce premier est correcte du fait qu'elle est extrapolée

du deuxième. En ayant comme outil de contrôle du DCDQ le MABC-2 et vice versa, il faut nuancer la validité concomitante de ces instruments car ils ne peuvent pas encore être considérés comme des gold standards.

On constate par ailleurs dans l'étude de Capistrano et al, 2015 des résultats disparates entre les différents évaluateurs, ici les parents, les professeurs et les enseignants d'éducation physique dans l'administration du MABC-2, outil non adapté pour eux car normalement réalisé par un professionnel qualifié. Les résultats sont malgré tout intéressants à analyser. Le professeur d'éducation physique semble plus capable que les autres et montre des meilleures compétences grâce à ses connaissances du mouvement supérieures à celles des parents et des professeurs. Il existe une différence significative entre le professeur d'éducation physique et les autres alors qu'il n'y en a pas entre les parents et les professeurs par le test post hoc de Tukey. Le kinésithérapeute étant le professionnel le plus qualifié concernant le mouvement sain et perturbé, on peut y voir une plus grande facilité de diagnostic que ces derniers pour le kinésithérapeute. Il est intéressant d'étudier sur un travail à grande échelle les résultats que l'on pourrait obtenir en comparaison à des professionnels qualifiés qui administrent traditionnellement le MABC-2 ou en comparaison avec les parents pour le DCDQ.

En s'intéressant de plus près à la construction de ce questionnaire et de ce test, on va pouvoir constater les corrélations entre chaque sous élément et le total afin de montrer la puissance d'association des items et du résultat final. Dans l'étude d'Ellinoudis et al, 2011, les trois composantes motrices de la batterie d'évaluation du mouvement chez l'enfant sont significativement corrélées mais modérément ce qui est optimal car les trois sous-échelles sont censées mesurer différents aspects du mouvement. En effet, si la corrélation était trop élevée, cela signifierait qu'elles mesurent les mêmes caractéristiques alors que si l'association était trop faible, c'est qu'elles mesurent des compétences non pertinentes. Le MABC-2 couvre un large éventail de capacités motrices en raison de l'hétérogénéité des TDC. Ainsi, la nature de cette pathologie englobe une multitude de déficits potentiels à évaluer. L'étude sur le DCDQ menée par Pannekoek en 2012, quant à elle, montre une corrélation entre les éléments de la même sous-échelle en moyenne supérieure que la corrélation entre les éléments de différentes sous-échelles. De la même façon que pour Ellinoudis et al, 2011, pour le MABC-2, on constate une corrélation positive et significative mais assez modérée pour être sur des caractéristiques différentes des TDC (corrélations entre le contrôle pendant le mouvement et la motricité fine/l'écriture manuscrite r = 0,52 ; entre le contrôle pendant le mouvement et la coordination générale r = 0,77 ; entre la motricité fine/l'écriture manuscrite et la coordination générale r = 0,73). Néanmoins, la corrélation entre chaque élément individuel et le questionnaire entier est importante (r = 0,62 à 0,82) donc ils mesurent une construction similaire, on remarque des corrélations positives entre r = 0,31 à r = 0,86 entre les éléments. De la même façon, les corrélations entre les sous-échelles sont semblables dans l'étude de Cairney en 2008 à celles montrées par Pannekoek et al, en 2012 (corrélations entre le contrôle pendant le mouvement et la motricité fine/l'écriture manuscrite r = 0,58 ; entre le contrôle pendant le mouvement et la coordination générale r = 0,75 ; entre la motricité fine/l'écriture manuscrite et la coordination générale r = 0,68) alors que les corrélations entre l'ensemble et les sous-échelles sont excellentes montrant une bonne construction, la pertinence et la performance de chaque sous-échelle (corrélations entre l'ensemble et le contrôle pendant le

mouvement r = 0,90 ; avec la motricité fine/l'écriture manuscrite r = 0,84 ; avec la coordination générale r = 0,92).

La cohérence interne est excellente pour le DCDQ et le MABC-2. On cherche à vérifier si les éléments du questionnaire mesurent la même construction, si l'alpha de Cronbach est supérieur à 0,80, la cohérence interne est bonne. Cette dernière est un indicateur de fiabilité. Dans l'étude de Martini et al, 2011, α = 0,95 pour le DCDQ'07 et α = 0,94 pour le DCDQ-FC. Chaque item du DCDQ'07 est corrélé avec le test global (r = 0,646 à 0,795) alors que chaque item du DCDQ-FC est corrélé avec le test global (r = 0,693 à 0,784). On peut retirer de cette étude que l'item « attraper une balle », « apprendre des nouvelles tâches motrices », « ranger, s'habiller, attacher ses chaussures » et « ne jamais être décrit comme un éléphant dans un magasin de porcelaine » sont les meilleures contributions au questionnaire. « Sauter par-dessus des obstacles facilement », « courir aussi rapidement que les autres enfants » et « écrire un dessin reconnaissable » sont les moins bons. L'étude dirigée par Ray-Kaeser en 2019 sur le DCDQ-FE montre une excellente cohérence interne α = 0,96. Quand chaque élément est supprimé, α = 0,95/0,96 ce qui montre qu'aucun élément ne pose un problème et chaque item mesure la construction. Ainsi la corrélation item-questionnaire global est r = 0,673 à 0,836. Wilson et al, 2009, ont eux aussi montré une excellente cohérence interne avec α = 0,94. En supprimant chaque item indépendamment on obtient également un alpha de Cronbach excellent (α > 0,94). La corrélation entre chaque item et le DCDQ total est entre r = 0,52 à 0,78. Pannekoek et al, 2012 en arrive à une conclusion semblable α = 0,95 pour le DCDQ total. La cohérence interne est très forte, chaque sous échelle a également un niveau très satisfaisant individuellement (le contrôle pendant le mouvement α = 0,94 ; la motricité fine/l'écriture manuscrite α = 0,88 ; la coordination générale α = 0,85). Tseng et al, 2010 étudiait un questionnaire à 17 articles (α = 0,84) alors que lorsqu'il est réduit à 15 articles en supprimant les performances de sport d'équipe et l'apprentissage des nouvelles compétences améliorant la cohérence interne α = 0,89. Cairney et al, 2008, donne les résultats de la cohérence interne du DCDQ et de ses sous-échelles et ils sont très bons (total α = 0,94 ; le contrôle pendant le mouvement α = 0,91 ; motricité fine α = 0,91 ; α = 0,84) au même titre que Caravale et al, 2015 dans de moins bonnes proportions (α = 0,89 ; α = 0,83 ; α = 0,79 ; α = 0,73). Ce même Caravale et al, en 2014 indique une excellente cohérence interne avec un alpha de Cronbach α = 0,94, entre 0,93 et 0,94 si un élément de la liste d'items est supprimé alors que chaque élément est corrélé avec le test de façon importante (r = 0,58 à 0,82). En ce qui concerne le MABC-2, la cohérence interne est également très bonne. Pour Schoemaker et al, 2012, α = 0,94 pour la liste de contrôle du MABC-2, la cohérence interne est très bonne et tous les éléments mesurent la même construction, chaque item aborde un aspect de la performance motrice fonctionnelle. Wuang et al, 2012 définit la cohérence interne du MABC-2 complet comme bonne α = 0,90 et celle des sous échelles comme telle également (dextérité manuelle α = 0,81 ; visée/capture α = 0,84 ; équilibre α = 0,88). On note une homogénéité suffisante de tous les domaines. Ellinoudis et al, 2011 donne les valeurs du coefficient alpha de Cronbach des sous-échelles du MABC-2, ils sont 0,51, 0,70 et 0,66 pour la dextérité manuelle, viser et attraper, et enfin l'équilibre, respectivement et elles sont inférieures aux normes recommandées.

La fiabilité test-retest a été jugée à plusieurs reprises pour le MABC-2 et le DCDQ. Premièrement, le DCDQ avec par exemple l'étude de Martini et al, 2011 qui montre des bonnes corrélations intra classes : intra langue 0,916 (test-retest dans la même langue avec deux administrateurs différents) ; inter langue 0,99 (premier testeur d'une langue et deuxième testeur dans l'autre) ; interlangue temporelle du français vers l'anglais 0,913 et de l'anglais vers le français 0,934. Les deux tests sont réalisés à une moyenne de 39 jours sans affecter la stabilité du questionnaire. Ray-Kaeser, en 2019, a montré de son côté dans une moyenne de 38 jours entre les deux tests, une corrélation intra classe de 0,865 avec un intervalle de confiance à 95% pour le groupe clinique et de 0,814 pour le groupe témoin donnant pour l'échantillon total un ICC de 0,956 ($p < 0,005$) avec le même intervalle de confiance. Le DCDQ-C a une corrélation de Pearson $r = 0,94$ avec $p < 0,001$ d'après l'étude de Tseng et al, 2012. Alors que pour Caravale et al, 2014 que 11 éléments du DCDQ-IT affichent un ICC entre 0,82 à 1, 3 éléments ont une ICC entre 0,62 à 0,75 alors que le dernier a un ICC de 0,43. Puis de son côté, le MABC-2 a également été étudié sur la fiabilité test-retest par Wuang et al, 2012 et Holm et al, 2013. Le premier a un ICC entre 0,88 à 0,99 pour les items et les sous-échelles alors que l'ICC est de 0,97 pour le score total. Le second indique un ICC de 0,23 à 0,76 avec un intervalle de confiance de 95%. La fiabilité test-retest du DCDQ est très bonne alors que celle du MABC-2 est encore approximative car pas assez étudiée et trop sensible aux variations individuelles de l'évaluateur, une première étude est excellente et une deuxième moins bonne.

La fiabilité inter évaluateurs a été évalué pour le MABC-2 seulement, il n'existe pas encore d'étude sur ce thème concernant le DCDQ. Jaikaew et al, 2019 donne pour valeur de fiabilité inter évaluateurs un intervalle de corrélations inter classes de 0,71 à 1 pour les éléments du test. Holm et al, 2013 constate lui des corrélations inter classes ICC = 0,35 à 0,67 avec un intervalle de confiance de 95%.

La fiabilité reste à développer de façon importante. La fiabilité inter évaluateurs (pour le DCDQ et le MABC-2) et test-retest (notamment du MABC-2) doivent encore être étudiées pour les deux outils diagnostics de cette revue.

Il est compliqué de parler de diagnostic des TDC avec un seul des deux instruments étudiés, vu les résultats précédemment décrits. On considère qu'ils sont de bons évaluateurs de difficultés motrices et en recoupant les deux instruments, c'est une excellente association pour le diagnostic. En effet, en évaluant deux critères différents, avec deux méthodes distinctes, le TDC peut plus facilement être découvert.

Il existe plusieurs risques de **biais** dans les études diagnostiques que nous allons évaluer pour cette revue sur le DCDQ et le MABC-2. [19]

L'intégrité des tests en est un, il est nécessaire que la connaissance du statut de la maladie (OUI/NON) des sujets n'influence pas l'évaluation du test (et vice versa), les opérateurs qui évaluent le test doivent être à l'insu du résultat du statut de la maladie (vice versa). Dans cette revue, on constate, à l'aide du tableau relatant la qualité méthodologique des études de fiabilité avec l'échelle QAREL, que cette évaluation en aveugle pour l'examinateur est toujours

respectée alors que dans les deux études utilisant un groupe clinique, le sujet évalué est conscient de son statut vis-à-vis de la pathologie.

Le biais de vérification est un deuxième risque de biais, dans les études, le test devrait être appliqué à tous les sujets mais on est parfois dans la situation où la référence n'est réalisée que si le test est positif pour contrôler éliminant de l'évaluation de vérification les négatifs dont les faux négatifs. Dans cette revue, ce biais est absent.

Il existe en revanche un biais sur le gold standard. En effet, il y a des **erreurs sur la référence**. Pour de nombreuses pathologies dont les TDC, il est impossible de déterminer avec certitude le statut de la maladie et les meilleures références peuvent donner des résultats faux et ainsi fausser la comparaison. En l'absence de preuves ou de résultats statistiques pas assez discriminants entre sain et atteint, il est compliqué de déterminer le potentiel d'un test par association. Notamment la sensibilité et la spécificité des tests considérées comme référence dans ce domaine qui peuvent être trop faibles. Ces erreurs peuvent avoir des conséquences sur l'évaluation d'un test diagnostic.

Le biais de recrutement apparaît quand les sujets malades ne sont pas représentatifs des sujets malades de la population ou quand les sujets contrôles (non malades) ne sont pas représentatifs des sujets non malades de la population, c'est-à-dire que les enfants TDC auraient de très mauvaises évaluations aux tests en présence (par exemple, sélection d'enfants uniquement présents dans la zone rouge du MABC-2 donc inférieur à 5% et pas dans la zone orange d'enfants suspectés entre 5 et 15%) alors que les enfants du groupe témoin auraient d'excellentes dispositions motrices entrainant des paramètres de validité des tests surestimés. Les estimations concernant la sensibilité et la spécificité ne sont alors pas extrapolables à la pratique courante pour lequel le test est développé, c'est le biais de sélection. Dans cette revue, seulement deux études se basent sur une comparaison groupe clinique et groupe contrôle mais leurs critères d'inclusion sont corrects et pas stratégiquement ou excessivement restrictifs dans le recrutement pouvant engendrer un spin positif.

Le biais d'interprétation intervient quand des informations extérieures (données cliniques ou résultats d'autres tests) peuvent influencer la procédure du test à évaluer qui ne sera pas appliquée ainsi dans la pratique ultérieure. Ce biais est également évalué et relaté dans le tableau récapitulatif de l'échelle méthodologique QAREL et les études respectent cet « aveuglement » des données cliniques non destinées à la procédure des tests.

Le biais d'extrapolation comprend plusieurs facteurs qui peuvent influencer les performances d'un test pour détecter une maladie. Il existe les facteurs liés à l'opérateur (son expérience, sa connaissance du test, …), les facteurs liés au sujet (caractéristiques géographiques, socioéconomiques, culturelles, …), l'environnement dans lequel le test est réalisé (cadre de l'évaluation, ressources disponibles, accès aux traitements, prévalence de la maladie, …). Ainsi, les performances d'un test réalisé dans une population ne peuvent pas être systématiquement extrapolées à d'autres populations (par exemple la différence entre le Brésil et les Etats-Unis en termes de prévalence et du niveau moteur global mais aussi la différence entre un échantillon cible et la population générale) ou encore les résultats d'un outil diagnostic ne peuvent pas toujours être transposés à un autre contexte (changement

d'évaluateur ou de cadre clinique). Il est ainsi difficile de présumer de la validité d'un test à partir d'un petit échantillon réalisé dans un environnement particulier et l'adaptation clinique reste trop souvent basée sur des conjectures plutôt que sur les preuves dans le domaine de l'enfant TDC.

Le biais d'incorporation survient lorsque le résultat du test est incorporé dans le diagnostic. En l'absence de gold standard, le test diagnostic est intégré dans la stratégie de référence. Dans l'évaluation des propriétés d'un test diagnostic, il existe ce biais d'incorporation lorsque le résultat du test diagnostic évalué est inclus dans le résultat donné par la méthode de référence. Cela survient notamment lorsque la méthode de référence repose sur plusieurs examens du fait de l'inexistence d'un examen de référence unique comme dans la détection des TDC et donc de cette revue.

La revue n'est ainsi pas à l'abri des risques de biais spécifiques à l'étude diagnostique mais ils sont limités au maximum de ce qui est possible par rapport à la littérature scientifique sur les TDC.

La méta-analyse a pour but de quantifier un effet commun à toutes les études incluses. Cela suppose une homogénéité c'est-à-dire qu'elles étudient bien toutes le même effet. Afin de tester cette hypothèse, on doit réaliser un test d'homogénéité afin de vérifier si au moins l'une de ces études ne se différencie pas de l'ensemble. Le sujet de la revue concerne la validité et la fiabilité d'un test diagnostic et possède plusieurs sous catégories, il y a ainsi plusieurs éléments à étudier. Pas toutes les études incluses dans la revue s'intéressent à tous les éléments mais à partir du moment où un des domaines est approfondi en respectant les critères de qualité d'une étude, elle revêt un caractère acceptable pour être sélectionnée dans la revue. La population, l'intervention et la méthode de mesure sont des caractéristiques qui demeurent semblables dans la sélection incluse malgré la diversité des sujets, on observe une absence d'hétérogénéité des études. [50]

4.2 Applicabilité des résultats en pratique clinique

Après analyse des différents articles, la question de son applicabilité dans la pratique clinique pour le bénéfice du patient, du professionnel de santé et pour une amélioration de la santé publique se pose. Il s'agit de démontrer et justifier ces instruments que sont le DCDQ et le MABC-2 en fournissant de la littérature scientifique de qualité sur sa validité et sa fiabilité pour le retranscrire dans une utilisation collective éclairée et efficace. L'intérêt est de percevoir un progrès dans le diagnostic de la pathologie en fournissant des données favorisant leur application ou encore en élargissant les praticiens qui s'en servent et d'intégrer le kinésithérapeute à la liste des administrateurs et ainsi offrir une option supplémentaire permettant alors de réduire les difficultés pour ces enfants en pouvant les orienter vers un thérapeute qualifié, améliorer la prise en charge et obtenir une meilleure santé de la population générale mais aussi de mieux connaitre la prévalence de la maladie en diminuant le pourcentage d'enfants non détectés. Il faut voir cette revue comme une opportunité d'intéresser les professionnels de santé quels qu'ils soient aux troubles moteurs, notamment ceux du développement ayant un impact important pendant l'enfance et se répercutant à la vie adulte sur tout un groupe de la population et montrer l'accessibilité et la performance de

ces outils diagnostics. Le dépistage des TDC revêt une utilité de santé publique et il s'agit alors de fournir des éléments de diagnostic validés, reproductibles, universels, simples, faciles à mettre en place (matériel requis) et réalisables en routine (test pas trop long) et peu coûteux. Dans le domaine de la santé, lorsqu'il s'agit de faire un diagnostic de masse, il est rarement possible de soumettre chaque sujet à un examen clinique détaillé et éventuellement à plusieurs examens complémentaires, ce qui représente une procédure longue et coûteuse. On doit donc se contenter le plus souvent de procédures plus simples et plus rapides, même si elles sont approximatives et associées à un certain taux d'erreur. L'objectif du dépistage est d'améliorer la santé des individus par le diagnostic précoce des maladies à un stade où elles sont curables ou quand leurs conséquences peuvent être limitées. Il est ainsi important que les qualités intrinsèques du test diagnostic soient très bonnes. En effet, il faut que le nombre de vrais positifs soit optimal pour que le plus de patients soient justement évalués et traités mais il faut aussi que le nombre de faux positifs soit le plus faible car ce sont des sujets non atteints qui vont engendrer des coûts de santé évitables en lançant le processus de soin (en plus du traumatisme psychologique) et les dépenses superflues doivent être annulées. Les faux négatifs vont également engendrer des coûts à cause des examens complémentaires avant de pouvoir poser le bon diagnostic. Dans une optique de dépistage, la VPP donne une idée du nombre d'examens inutiles qui seront faits à une population en bonne santé. Il faut que le coût de la recherche des cas ne soit pas disproportionné par rapport au coût global des soins médicaux, en cela le DCDQ et le MABC-2 sont excellents de par leur composition et leur administration. En effet, le DCDQ est un questionnaire d'observations à faire en une quinzaine de minutes et le MABC-2 est une révision du MABC [51], trop long et compliqué à mettre en application, même si sa validité et sa fiabilité ont été démontrés, les professionnels ont préféré renouveler le modèle pour plus de simplicité et d'efficacité et pour s'adapter aux normes récentes. Elle se compose de deux parties, le test de performance qui s'adapte en fonction du patient (l'âge notamment) et une liste à cocher. Le temps de passation du test est entre 20 à 40 min et le questionnaire environ 10 min. Ces instruments diagnostics sont assez exhaustifs et représentatifs de l'ensemble du spectre des TDC.

Les études incluses portent malheureusement très souvent sur un échantillon réduit, il est difficile d'organiser une étude de grande échelle qualitative et pertinente. Les TDC concernent autour de 4 millions de français, c'est un groupe clinique potentiel très conséquent et définir une stratégie diagnostique efficace est essentiel.

4.3 Qualité des preuves

Le nombre conséquent d'essais incorporés dans la revue représente un avantage méthodologique et permet d'arriver à une conclusion solide sur les qualités diagnostiques du DCDQ et du MABC-2. Les études présentent des protocoles protéiformes s'intéressant à la dimension hétérogène du diagnostic des TDC. La qualité des preuves en ce qui concerne les études diagnostiques, par leur structure et leur construction méthodologique intrinsèque, sont souvent inférieures à la norme. En effet, il existe un faible niveau de preuve scientifique pour les études épidémiologiques descriptives transversales et elles sont catégorisés niveau 4 et elles sont au grade C de recommandations selon l'état des lieux de l'HAS de 2013 (tableau extrait du document "État des lieux : Le niveau de preuve et gradation des recommandations

de bonne pratique" de l'HAS en 2013). Cette gradation des recommandations fondée sur le niveau de preuve scientifique de la littérature venant à l'appui de ces recommandations ne présume pas obligatoirement du degré de force de ces recommandations. Cependant, il peut exister des recommandations de grade C ou fondées sur un accord d'experts néanmoins fortes malgré l'absence d'un appui scientifique. Les raisons de cette absence de données scientifiques peuvent être multiples. Il est donc utile de préciser la relation à laquelle on doit s'attendre entre gradation et hiérarchisation des recommandations. [18]

D'après la publication de l'HAS et son service des bonnes pratiques professionnelles d'avril 2013, le groupe de travail GRADE (The Grading of Recommendations Assessment, Development and Evaluation) a débuté en 2000 comme une collaboration informelle de personnes intéressées par l'évaluation des défauts des systèmes de gradation actuels dans les soins de santé. La description du système GRADE a été rédigée à partir des articles à ce sujet publiés en 2004, 2008 et 2010 [52]. Le système GRADE est centré sur les résultats et est devenu une référence dans la littérature scientifique. Toute question concernant une prise en charge clinique comporte quatre composants : la population de patients, l'intervention d'intérêt, le comparateur, et les résultats d'intérêt (PICO).

On remarque en fonction des facteurs influençant les données scientifiques que les études présentées dans cette étude sont réalisées avec une grande rigueur avec des risques de biais diminués au maximum et les études incluses sont les plus pertinentes de la littérature actuelle avec un type d'études qui est le meilleur disponible mais qui n'est pas optimal car aucun gold standard n'existe. Les résultats montrent encore une certaine hétérogénéité dans l'évaluation de la validité prédictive et sont basés sur des études à faible échantillon. Ainsi, on ne peut pas considérer cette revue comme avec une qualité de preuves élevée mais plutôt faible ou modérée car les professionnels sont tout de même confiants par rapport aux conclusions et à l'estimation de l'effet. De ce fait, les outils diagnostics évalués sont très largement utilisés dans la pratique clinique quotidienne.

Selon une autre grille d'évaluation jugeant du niveau de preuve des études diagnostiques du Center for Evidence Based Medicine (CEBM) d'Oxford, le niveau de preuve qui correspond à cette revue est le niveau 4 car il n'existe pas standard de référence indépendant et validé. Structurellement, une étude sur les TDC ne peut pour l'instant être considérée comme démontrée scientifiquement par la littérature scientifique selon cette grille. [53]

Tableau 2 Niveau de preuve des études diagnostiques, adaptée de la grille du CEBM Oxford (Source : www.cebm.net).
Table 2 Level of evidence for diagnostic studies, adapted from the CEBM "Levels of Evidence" document (http// :www.cebm.net).

Méta-analyse d'études de cohortes de validation Étude de cohortes de validation avec de bons standards de référence	**Niveau de preuve 1**
Méta-analyse d'études de cohortes exploratoires Étude de cohortes exploratoires avec de bons standards de référence	**Niveau de preuve 2**
Étude de sujets non consécutifs ou sans application homogène du/des standards de référence	**Niveau de preuve 3**
Étude cas-témoins avec mauvais ou sans standard de référence indépendant	**Niveau de preuve 4**

4.4 Biais potentiels de la revue

L'échelle de qualité méthodologique AMSTAR-2, sous forme d'une fiche synthétique, est utile pour coter la validité interne de revues de littérature systématiques. Les revues de la littérature sont des études secondaires : elles synthétisent des informations à partir d'études primaires. Lorsqu'elles présentent une méthode d'élaboration, elles sont qualifiées de systématiques. L'échelle de qualité méthodologique pour coter la validité interne de ce type d'étude est l'échelle AMSTAR-2. On choisit alors de s'appuyer sur cette grille pour faire l'analyse critique de notre revue systématique. Pour ce travail de recherche quelques limites ont été rencontrées et certains biais ont pu impacter la qualité de la revue. [54]

Premièrement, ce travail de recherche a été réalisé par une seule personne que ce soit pour la validation de la qualité méthodologique des études, l'extraction des données et l'interprétation des résultats, des biais potentiels auraient pu, peut-être, être évités si le travail avait été effectué collectivement.

Ensuite, une limite au cours de la recherche bibliographique est que les études non publiées en anglais ont été exclues de la revue notamment une étude chinoise de grande ampleur régulièrement recommandée dans les études incluses entrainant une perte d'information pertinente et constitue un biais linguistique. La revue explicite clairement les biais potentiels d'une étude diagnostique, évalue les biais présents dans les études sélectionnées dans la discussion en les prenant en compte. Les échelles méthodologiques pour juger l'inclusion d'un article définissent également les risques de biais présents dans chaque étude. L'auteur discute également de leur impact potentiel sur la synthèse de résultats de l'étude individuellement et de la méta analyse. La combinaison statistique des résultats de la revue utilise la méthode la plus appropriée possible, l'hétérogénéité de la nature des données obtenues et la variabilité des interventions rendent difficile le regroupement. Il existe une certaine hétérogénéité dans les résultats de la revue, notamment de la validité prédictive, qui est expliquée correctement et débattue. Les possibles biais de publication ont été évités en intégrant des études avec des résultats inégaux et contrastés dans la revue. Il ne s'agissait pas d'obtenir une validité du test ou du questionnaire si elle n'existe pas car l'objectif est d'augmenter le nombre d'enfants détectés justement et pas d'avoir des instruments incertains quant à leur efficacité. Les questions de recherche et les critères d'inclusion ont inclus les critères PICO. La méthode a été définie avant de réaliser la revue et il n'y a pas eu de déviation significative. De la même façon, le schéma d'étude à inclure dans la revue systématique a été expliqué dans la méthodologie. De plus, la revue peut être considérée comme exhaustive. En effet, plusieurs sources électroniques de recherche ont été utilisées et une stratégie complémentaire de recherche dans la littérature grise a été appliquée. Les études exclues ont été justifiées dans un tableau récapitulatif, de même que les incluses ont été détaillées dans un tableau synthétique en plus de l'évaluation méthodologique des études grâce aux échelles QAREL et QUADAS 2.

La seule source potentielle de conflit d'intérêts de cette revue systématique sur les TDC est que l'auteur est un enfant diagnostiqué avec des TDC. Cela pourrait engendrer un spin positif par une volonté inconsciente de posséder un arsenal diagnostic de qualité pour traiter ensuite de manière plus efficace et ciblée et afin que le plus grand nombre de patients suivent une

thérapie adéquate. L'étude n'est pas financée et l'auteur n'est pas acteur du système en place actuellement pouvant entraîner des intérêts ou des conflits. Chaque étude prise indépendamment a déclaré les conflits d'intérêts ou le financement du travail.

Une faiblesse de cette revue de littérature est le caractère scientifiquement indémontré du diagnostic des TDC qui restent eux-mêmes encore parfois opaques sur leurs origines, encore à l'étude. Les données actuelles de la science sont encore à affiner et à valider. L'avantage ici se porte sur le fait que la recherche a été assez centrée malgré le profil hautement hétérogène de la pathologie et des outils diagnostics à l'étude. L'utilisation des mots MeSH dans l'équation a permis un résultat de recherche non subjectif et reproductible.

Des doutes subsistent malgré tout sur la qualité de la revue en raison du faible niveau de preuves des études sélectionnées. Cela peut représenter un biais important dans la prise en compte des résultats pour chacune de ces études et interroger sur la véracité de notre conclusion. Devant l'impossibilité de trouver des études donnant un niveau de preuves supérieur, les informations données restent intéressantes à utiliser pour les professionnels de santé et peuvent avoir un intérêt dans la littérature scientifique des TDC.

5 Conclusion

L'objectif de cette étude était de déterminer la validité et la fiabilité de deux outils diagnostics, le DCDQ et le MABC-2, pour offrir des instruments de qualité aux praticiens permettant d'améliorer le pouvoir diagnostic des TDC et la prise en charge des troubles moteurs dans une population d'enfants. Il semble, à la vue de cette revue, difficile d'affirmer qu'il est possible de dépister avec une assez grande certitude les TDC avec l'un ou l'autre des outils simplement dans une population globale d'enfants. Cependant, sa capacité à déterminer un enfant malade d'un non atteint est correcte dans une population suspectée de troubles moteurs. La force diagnostique du MABC-2 et du DCDQ est plutôt bonne dans cette situation, c'est-à-dire celle présente dans la réalité clinique du quotidien. Les résultats du MABC-2 montrent une excellente cohérence interne ainsi qu'une bonne fiabilité statistique test-retest et inter-évaluateurs mais une validité prédictive contrastée même si la validité concomitante par rapport aux comparateurs est bonne. Or, on sait qu'il est déjà très largement utilisé seul par les examinateurs dans la pratique clinique malgré l'absence de données scientifiques offrant des preuves de la qualité du test. Il comporte cependant de nombreuses informations importantes et exploitables en étant mis en corrélation avec d'autres s'intéressant à un autre aspect de la maladie. Les résultats du DCDQ montrent lui aussi une cohérence interne ainsi qu'une bonne fiabilité test-retest mais manque d'études concernant la fiabilité inter-évaluateurs. La validité prédictive du questionnaire est meilleure que celle du MABC-2 mais est encore inférieure à la norme statistique demandée. Le DCDQ est quand même significativement corrélé avec les standards référencés.

On remarque que ces deux outils représentent assez bien l'hétérogénéité des TDC en évaluant des aspects différents de la pathologie tout en étant corrélés de façon significative et en montrant une validité simultanée, cela offre une pertinente homogénéité de l'hypothèse diagnostic. La combinaison des deux semble être la meilleure association pour obtenir un diagnostic efficace en concertation entre les professionnels de santé et l'entourage de l'enfant. Ces outils peuvent être appliqués par un kinésithérapeute formé à l'administration de l'un ou de l'autre. Celui-ci bénéficierait alors de son expertise dans l'analyse du mouvement sain ou perturbé et permettrait d'améliorer l'efficacité de l'évaluation des troubles moteurs dans la population mais aussi une progression de la santé publique avec un meilleur traitement du groupe atteint entrainant une augmentation du bien-être de la population. On constaterait une diminution des coûts que représentent les troubles engendrés par la pathologie dans la société civile mais aussi par les examens supplémentaires nécessaires pour définir la maladie du patient. Assimiler les deux instruments à une administration parallèle présente de nombreux avantages et aboutit à une meilleure capacité diagnostique, il est donc intéressant de développer la littérature sur le sujet en comparant les résultats statistiques entre une administration commune du MABC-2 et du DCDQ, une administration seule du MABC-2 et une administration seule du DCDQ. Modifier l'administration du DCDQ en le proposant à des professionnels qualifiés au contact régulier du patient plutôt qu'aux parents est également une option à développer dans la recherche scientifique. Tout ceci conduit alors à une pratique clinique de meilleure qualité et à un développement de la littérature sur un

domaine de la santé qui demande à être perfectionner et à fournir des preuves et des recommandations éclairées.

6 Bibliographie

[1] Trouble développemental de la coordination ou dyspraxie Synthèse et recommandations EXPERTISE COLLECTIVE. Inserm 2019.

[2] Trouble développemental de la coordination n.d. https://www.chusj.org/fr/soins-services/T/TDC.

[3] Dyspraxie ? TAC ? TDC ? Classification des troubles. https://www.dyspraxie.info/2015-04-14-17-30-00/dyspraxie-tac-tdc-un-peu-de-terminologie/313-dyspraxie-tac-tdc-classification-des-troubles.

[4] Blank R, Barnett AL, John Cairney |, Dido Green |, Kirby A, Polatajko H, et al. DEVELOPMENTAL MEDICINE & CHILD NEUROLOGY EACD RECOMMENDATIONS International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder 2019. https://doi.org/10.1111/dmcn.14132#support-information-section.

[5] Le point sur le trouble développemental de la coordination ou TDC - Institut TA n.d. https://institutta.com/mediatheque/trouble-developpemental-coordination-tdc.

[6] Magnat J, Xavier J, Zammouri I, Cohen D. Troubles développementaux de la coordination (TDC) : perspective clinique et synthèse de l'état des connaissances. Neuropsychiatr Enfance Adolesc 2015;63:446–56. https://doi.org/10.1016/j.neurenf.2015.09.004.

[7] Les dyspraxies | Dyspraxie France Dys n.d. https://www.dyspraxies.fr/les-dyspraxies/les-dyspraxies-et-autres-troubles-dys-presentation-et-informations-generalistes/.

[8] TDC ou la dyspraxie, académie de Bordeaux, Madame Kozakevich, https://blogacabdx.ac-bordeaux.fr/ash40/wp-content/uploads/sites/85/2018/05/DYSPRAXIE30052018.pdf

[9] Albaret J-M. Troubles psychomoteurs et obésité pédiatrique : Quels liens ?). http://www.cnreppop.com/pdf/Albaret-mai2017.pdf.

[10] Henderson SE, Sugden DA, Barnett A. Fiche descriptive-Outil d'évaluation pédiatrique Portail Enfance et Familles MABC-2-Movement Assessment Battery for Children-Second Edition Fiche descriptive Outil d'évaluation pédiatrique Instrument MABC-2-Movement Assessment Battery for Children, Second Edition par. 2007.

[11] MABC-2 - Batterie d'évaluation du mouvement chez l'enfant - 2nde édition - Pearson Clinical & Talent Assessment n.d. https://www.pearsonclinical.fr/mabc-2-batterie-d-evaluation-du-mouvement-chez-l-enfant-seconde-edition.

[12] MABC-2 BATTERIE D'EVALUATION DU MOUVEMENT CHEZ L'ENFANT-SECONDE. EDITION 2016, ADAPTATION FRANÇAISE : J. MARQUET-DOLÉAC, R. SOPPELSA, J.-M. ALBARET.

[13] THE DEVELOPMENTAL COORDINATION DISORDER QUESTIONNAIRE 2007. B. N. Wilson, 2007 www.dcdq.ca Wilson, B.N., Crawford, S.G., Green, D., Roberts, G., Aylott, A., &

Kaplan, B. (2009). Psychometric Properties of the Revised Developmental Coordination Disorder Questionnaire. Physical & Occupational Therapy in Pediatrics, 29(2):182-202..

[14] Ray-Kaeser D-F, Thommen S, Martini E, Jover R, Gurtner M, Bertrand B. QUESTIONNAIRE D'IDENTIFICATION DU TROUBLE DÉVELOPPEMENTAL DE LA COORDINATION (5-15 ANS) 2019. https://doi.org/10.1371/journal.pone.0217280.

[15] Généralités : Commission d'experts sur le TAC. MADIEU Emmanuel, psychomotricien ISRP Marseille 2018.

[16] KINÉSITHÉRAPIE : Etymologie de KINÉSITHÉRAPIE. https://www.cnrtl.fr/etymologie/kinésithérapie .

[17] Plans d'étude - CEBM n.d. https://www.cebm.net/2014/04/study-designs/.

[18] Niveau de preuve et gradation des recommandations de bonne pratique, HAS. 2013.

[19] Principes des études diagnostiques Caroline Elie URC Paris Centre Hôpital Necker. https://www.rechercheclinique pariscentre.fr/wp-content/uploads/2012/12/Diagnostic-DIU-2012_elie.pdf.

[20] Sensibilité, spécificité, valeur prédictive positive et valeur prédictive négative d'un test diagnostique - EM|consulte n.d. https://www.em-consulte.com/rmr/article/144227.

[21] Degoul S. UNIVERSITÉ DU DROIT ET DE LA SANTÉ - LILLE 2 FACULTÉ DE MÉDECINE HENRI WAREMBOURG, THÈSE POUR LE DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE Comment mesurer la performance d'un test diagnostique ? Présentation et comparaison d'indicateurs.

[22] Cohen J, Bossuyt P, Levy C, Chalumeau M, Martinot A, Launay E. Comment utiliser la grille STARD et l'outil QUADAS-2 ? vol. 22. 2015.

[23] Description of the QUADAS-2 critical appraisal checklist. https://journals.plos.org/.

[24] Rostagno S, Pallot A. QUADAS-2 : traduction française de l'échelle de qualité méthodologique pour les études diagnostiques de validité. Kinésithérapie, La Rev 2020. https://doi.org/10.1016/j.kine.2019.12.049.

[25] Pallot A, Rostagno S. QAREL : traduction française de l'échelle de qualité méthodologique pour les études diagnostiques de fiabilité. Kinesitherapie 2020. https://doi.org/10.1016/j.kine.2019.12.048.

[26] Modified Quality Appraisal of Diagnostic Reliability (QAREL) Checklist: Reliability Studies. https://www.ifec.net/download/Annexe4.pdf.

[27] Sarraff TDF de S, Martinez CMS, Santos JLF. Specificity and sensitivity of the DCDQ for children aged 8 to 10 years in Brazil. Rev Ter Ocup Da Univ São Paulo 2018;29:135–43. https://doi.org/10.11606/issn.2238-6149.v29i2p135-143.

[28] Is Questionnaire-Based Screening Part of the Solution to Waiting Lists for Children with Developmental Coordination Disorder ? . Dido Green, Tracie Bishop, Brenda N Wilson, Susan Crawford, Richard Hooper, Bonnie Kaplan and Gillian Baird.

[29] Kokštejn J, Musálek M, Tufano JJ. Construct validity of the movement assessment battery for children-second edition test in preschool children with respect to age and

gender. Front Pediatr 2018;6. https://doi.org/10.3389/fped.2018.00012.

[30] Psotta R, Brom O. Factorial Structure of the Movement Assessment Battery for Children Test—Second Edition in Preschool Children. Percept Mot Skills 2016;123:702–16. https://doi.org/10.1177/0031512516666072.

[31] The Developmental Coordination Disorder Questionnaire and Movement Assessment Battery for Children as a Diagnostic Method in Australian Children. Lauren R. Civetta, BPhty (Hons), GradDipPubHlth, and Susan L. Hillier, PhD, BAppSc (Physiotherapy) School of Health Sciences, University of South Australia, Adelaide, South Australia, Australia.

[32] Ferreira L, Gabbard C, Vieira JLL, da Silva PN, Cheuczuk F, da Rocha FF, et al. Reconsidering the use of cut-off scores: DCDQ-Brazil. Rev Bras Med Do Esporte 2019;25:344–8. https://doi.org/10.1590/1517-869220192504183194.

[33] Martini R, St-Pierre MF, Wilson BN. French Canadian cross-cultural adaptation of the Developmental Coordination Disorder Questionnaire'07: DCDQ-FC. Can J Occup Ther 2011;78:318–27. https://doi.org/10.2182/cjot.2011.78.5.7.

[34] Montoro APPN, Capistrano R, Ferrari EP, Da Silva Reis M, Cardoso FL, Beltrame TS. Concurrent validation of the MABC-2 and developmental coordination disorder questionnaire-BR. J Hum Growth Dev 2016;26:74–80. https://doi.org/10.7322/jhgd.110421.

[35] Capistrano R, Ferrari EP, De Souza LP, Beltrame TS, Cardoso FL. Concurrent validation of the MABC-2 motor tests and MABC-2 checklist according to the developmental coordination disorder questionnaire-BR. Motriz Rev Educ Fis 2015;21:100–6. https://doi.org/10.1590/S1980-65742015000100013.

[36] Schoemaker MM, Niemeijer AS, Flapper BCT, Smits-Engelsman BCM. Validity and reliability of the Movement Assessment Battery for Children-2 Checklist for children with and without motor impairments. Dev Med Child Neurol 2012;54:368–75. https://doi.org/10.1111/j.1469-8749.2012.04226.x.

[37] Asunta P, Viholainen H, Ahonen T, Rintala P. Psychometric properties of observational tools for identifying motor difficulties - A systematic review. BMC Pediatr 2019;19. https://doi.org/10.1186/s12887-019-1657-6.

[38] Ray-Kaeser S, Thommen E, Martini R, Jover M, Gurtner B, Bertrand AM. Psychometric assessment of the French European Developmental Coordination Disorder Questionnaire (DCDQ-FE). PLoS One 2019;14. https://doi.org/10.1371/journal.pone.0217280.

[39] Wilson BN, Crawford SG, Green D, Roberts G, Aylott A, Kaplan BJ. Psychometric properties of the revised developmental coordination disorder questionnaire. Phys Occup Ther Pediatr 2009;29:182–202. https://doi.org/10.1080/01942630902784761.

[40] Jaikaew R, Satiansukpong N. Movement assessment battery for children-second edition (MABC2): Cross-cultural validity, content validity, and interrater reliability in Thai children. Occup Ther Int 2019;2019. https://doi.org/10.1155/2019/4086594.

[41] Ellinoudis T, Evaggelinou C, Kourtessis T, Konstantinidou Z, Venetsanou F, Kambas A.

Reliability and validity of age band 1 of the Movement Assessment Battery for Children - Second Edition. Res Dev Disabil 2011;32:1046–51. https://doi.org/10.1016/j.ridd.2011.01.035.

[42] Wuang YP, Su JH, Su CY. Reliability and responsiveness of the Movement Assessment Battery for Children-Second Edition Test in children with developmental coordination disorder. Dev Med Child Neurol 2012;54:160–5. https://doi.org/10.1111/j.1469-8749.2011.04177.x.

[43] Pannekoek L, Rigoli D, Piek J, Barrett N, Piek JP, Barrett NC, et al. The Revised DCDQ: Is It a Suitable Screening Measure for Motor Difficulties in Adolescents? vol. 29. 2012.

[44] Tseng MH, Fu CP, Wilson BN, Hu FC. Psychometric properties of a Chinese version of the Developmental Coordination Disorder Questionnaire in community-based children. Res Dev Disabil 2010;31:33–45. https://doi.org/10.1016/j.ridd.2009.07.018.

[45] Rivard L, Missiuna C, Mccauley D, Cairney J. Descriptive and factor analysis of the Developmental Coordination Disorder Questionnaire (DCDQ'07) in a population-based sample of children with and without Developmental Coordination Disorder. Child Care Health Dev 2014;40:42–9. https://doi.org/10.1111/j.1365-2214.2012.01425.x.

[46] Cairney J, Missiuna C, Veldhuizen S, Wilson B. Evaluation of the psychometric properties of the developmental coordination disorder questionnaire for parents (DCD-Q): Results from a community based study of school-aged children. Hum Mov Sci 2008;27:932–40. https://doi.org/10.1016/j.humov.2008.04.002.

[47] Caravale B, Baldi S, Capone L, Presaghi F, Balottin U, Zoppello M. Psychometric properties of the Italian version of the Developmental Coordination Disorder Questionnaire (DCDQ-Italian). Res Dev Disabil 2015;36:543–50. https://doi.org/10.1016/j.ridd.2014.10.035.

[48] Caravale B, Baldi S, Gasparini C, Wilson BN. Cross-cultural adaptation, reliability and predictive validity of the Italian version of Developmental Coordination Disorder Questionnaire (DCDQ). Eur J Paediatr Neurol 2014;18:267–72. https://doi.org/10.1016/j.ejpn.2013.11.009.

[49] Holm I, Tveter AT, Aulie VS, Stuge B. High intra- and inter-rater chance variation of the movement assessment battery for children 2, ageband 2. Res Dev Disabil 2013;34:795–800. https://doi.org/10.1016/j.ridd.2012.11.002.

[50] Maison P. La méta-analyse sur données résumées. Rech Soins Infirm 2010:18–24. https://doi.org/10.3917/rsi.101.0018.

[51] Henderson SE, Sugden DA, Soppelsa R, Albaret J-M, Valencia Berclaz S. Batterie d'Evaluation du Mouvement chez l'Enfant. 2011.

[52] Méthodologie de gradation des recommandations en pratique vaccinale. Haut conseil de la santé publique, 12 janvier 2016. https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=538

[53] Howick J, Chalmers I, Lind J, Glasziou P, Greenhalgh T, Heneghan C, et al. Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence.

[54] Pallot A, Rostagno S. AMSTAR-2 : traduction française de l'échelle de qualité méthodologique pour les revues de littérature systématiques. Kinésithérapie, La Rev 2020. https://doi.org/10.1016/j.kine.2019.12.050.

Annexes

Annexe 1 : Le système GRADE

Une question implique souvent une autre précision : le contexte des soins dans lequel la recommandation sera mise en œuvre. Les auteurs incitent ceux qui élaborent des recommandations à préciser, en début de projet, tous les résultats importants par rapport au patient (bénéfices, inconvénients et coûts) et à distinguer parmi ceux-ci, les résultats décisifs. Ils proposent l'utilisation d'une échelle de 1 à 9 pour juger de l'importance des résultats. (7 – 9 : résultats décisifs ; 4 – 6 : résultats importants mais non décisifs ; 1 – 3 : résultats d'importance limitée). Pour une revue systématique, la qualité des données scientifiques reflète notre confiance dans le fait qu'une estimation de l'effet est correcte.

La qualité des données scientifiques pour chaque résultat important peut être déterminée après avoir considéré le type d'études, la qualité des études, l'homogénéité des résultats, le caractère direct des données scientifiques. Pour déterminer la qualité des données scientifiques, le système GRADE part du type d'étude. Il classe initialement les données en se fondant sur le type d'étude dont elles sont issues. Il distingue deux catégories : les essais contrôlés randomisés qui fournissent généralement des données scientifiques de qualité élevée et les études observationnelles qui fournissent généralement des données scientifiques de qualité faible. Puis il s'agit de considérer si les études ont des limites sérieuses, s'il y a une hétérogénéité importante des résultats, et si des doutes sur le caractère direct des données sont justifiés. Il s'agit de la qualité des données scientifiques pour chaque résultat important dans toutes les études. Cela ne signifie pas évaluer le niveau de chaque étude individuellement.
Ce système GRADE est intéressant dans le cadre de cette revue car il est possible de l'utiliser pour les tests ou les stratégies diagnostiques. Un groupe de travail considérant un test ou une stratégie diagnostique devrait commencer par identifier les patients, le test ou la stratégie diagnostique, le comparateur, et les résultats d'intérêt. En termes de résultats importants pour le patient, la validité diagnostique d'un test est un résultat intermédiaire ou de substitution. La meilleure manière d'évaluer une stratégie diagnostique est un essai contrôlé randomisé, dans lequel les investigateurs randomisent les patients en deux groupes (approche diagnostique expérimentale ou contrôle). Dans cette revue, pour un sujet comme l'évaluation des enfants suspectés de TDC, on ne s'intéresse pas à la thérapie après le test mais à la qualité de ces instruments. Le groupe de travail doit se centrer sur des études portant sur la validité diagnostique des tests. Les questions clés sont : y aura-t-il une réduction des faux négatifs ou des faux positifs et une augmentation correspondante des vrais positifs et des vrais négatifs ? Des patients semblables ou différents sont-ils classés de manière fiable par les stratégies diagnostiques alternatives ? Quels sont les résultats chez les patients classés comme cas et ceux classés comme n'ayant pas la maladie ?

Il existe quatre niveaux pour juger de la qualité des données scientifiques à l'appui des recommandations représentant un gradient de confiance dans l'estimation de l'effet d'un test ou d'une stratégie diagnostique sur des résultats importants pour le patient. Des études portant sur la validité diagnostique des tests sont classées initialement dans la catégorie

qualité élevée. Cependant, beaucoup de ces études apportent finalement des données scientifiques de qualité faible pour étayer les recommandations, car il s'agit de données scientifiques indirectes de l'impact du test sur les résultats importants pour le patient.

Niveaux de qualité des données scientifiques pour chaque résultat important d'après Balshem et al., 2011

Niveau de qualité	Définition*
Élevé	**Nous avons une confiance élevée dans l'estimation de l'effet : celle-ci doit être très proche du véritable effet.**
Modéré	Nous avons une confiance modérée dans l'estimation de l'effet : celle-ci est probablement proche du véritable effet, mais il est possible qu'elle soit nettement différente.
Faible	**Nous avons une confiance limitée dans l'estimation de l'effet : celle-ci peut être nettement différente du véritable effet.**
Très faible	Nous avons très peu confiance dans l'estimation de l'effet : il est probable que celle-ci soit nettement différente du véritable effet.

Facteurs qui influencent la qualité des données scientifiques des études sur la validité diagnostique des tests d'après le groupe de travail GRADE, 2008

<table>
<tr><td>Facteur</td><td>Explication</td></tr>
<tr><td>Type d'études</td><td>Les données scientifiques issues d'études transversales ou les études de cohortes, incluant des patients chez lesquels il existe des incertitudes diagnostiques (i.e. : patients auxquels les cliniciens appliqueraient le test en pratique courante), impliquant une comparaison directe entre le test considéré et un test de référence adapté (gold standard) sont considérées de qualité élevée.</td></tr>
<tr><td>Risques de biais (limites)</td><td>Des patients consécutifs devraient être recrutés comme une seule cohorte et non classés par stade de la maladie, et la manière dont les patients ont été adressés ainsi que leur sélection doivent être clairement décrits. Les tests devraient être réalisés chez tous les patients d'une même population de patients pour le nouveau test et le test de référence, par des investigateurs en insu des résultats de l'autre test.</td></tr>
<tr><td>Caractère indirect des données scientifiques</td><td></td></tr>
<tr><td>- résultats</td><td>Il existe souvent une absence de données scientifiques directes de l'impact du test sur des résultats importants pour le patient.
Le groupe de travail doit faire des déductions à partir des études portant sur la validité des tests pour ce qui est du rapport entre :
- d'une part les influences présumées des différences dans les VP, FP, VN, FN sur résultats importants pour le patient ;
- et d'autre part les complications et les coûts du test.
<table>
<tr><td colspan="3">Exemple de résultats d'un test et son impact attendu sur la prise en charge thérapeutique</td></tr>
<tr><td></td><td>Malades</td><td>Non malades</td></tr>
<tr><td>Test positif</td><td>Vrais positifs</td><td>Faux positifs</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td></td><td>
<table>
<tr><td></td><td>Administration d'un traitement d'efficacité connue</td><td>Probable anxiété et possible morbidité liée aux examens supplémentaire et au traitement</td></tr>
<tr><td>Test négatif</td><td>Faux négatifs
Conséquences possibles du retard de diagnostic (ne reçoit pas les traitements disponibles efficaces)</td><td>Vrais négatifs
Epargne au patient des effets secondaires potentiels du test de référence, patient rassuré.</td></tr>
</table>
C'est pourquoi le plus souvent, les études portant sur la validité diagnostique des tests n'apportent que des données scientifiques de qualité faible pour faire des recommandations en raison du caractère indirect des résultats (sur des critères de jugement intermédiaire ou de substitution).
</td></tr>
<tr><td>- populations, test diagnostic, test de comparaison et comparaisons indirectes</td><td>La qualité des données scientifiques peut être diminuée :
- s'il existe des différences importantes : entre la population étudiée et la population cible de la recommandation ; dans les tests étudiés et dans l'expertise diagnostique de ceux qui réalisent les tests dans les études comparé à ceux qui les réalisent dans l'environnement clinique cible de la recommandation ;
- si chacun des tests examinés est comparé au même test de référence dans des études différentes, et ne sont pas comparés directement dans les mêmes études.</td></tr>
<tr><td>Hétérogénéité importante des résultats</td><td>Une hétérogénéité inexpliquée de la sensibilité et de la spécificité peut diminuer la qualité des données scientifiques.</td></tr>
<tr><td>Probabilité élevée de biais de publication</td><td>Données scientifiques issues de petites études pour des tests nouveaux ou graphique en funnel plot.</td></tr>
</table>

Annexe 2 : Echelle AMSTAR-2

Tableau I. Traduction française de AMSTAR-2.

Items	Cotation
1	Est-ce que les questions de recherche et les critères d'inclusion de la revue ont inclus les critères PICO ?
2	Est-ce que le rapport de la revue contenait une déclaration explicite indiquant que la méthode de la revue a été établie avant de conduire la revue ? Est-ce que le rapport justifiait toute déviation significative par rapport au protocole ?
3	Les auteurs ont-ils expliqué leur choix de schémas d'étude à inclure dans la revue ?
4	Les auteurs ont-ils utilisé une stratégie de recherche documentaire exhaustive ?
5	Les auteurs ont-ils effectué en double la sélection des études ?
6	Les auteurs ont-ils effectuée en double l'extraction des données ?
7	Les auteurs ont-ils fourni une liste des études exclues et justifié les exclusions ?
8	Les auteurs ont-ils décrit les études incluses de manière suffisamment détaillée ?
9	Les auteurs ont-ils utilisé une technique satisfaisante pour évaluer le risque de biais des études individuelles incluses dans la revue ?
10	Les auteurs ont-ils indiqué les sources de financement des études incluses dans la revue ?
11	Si une méta-analyse a été effectuée, les auteurs ont-ils utilisé des méthodes appropriées pour la combinaison statistique des résultats ?
12	Si une méta-analyse a été effectuée, les auteurs ont-ils évalué l'impact potentiel des risques de biais des études individuelles sur les résultats de la méta-analyse ou d'une autre synthèse des preuves ?
13	Les auteurs ont-ils pris en compte le risque de biais des études individuelles lors de l'interprétation / de la discussion des résultats de la revue ?
14	Les auteurs ont-ils fourni une explication satisfaisante pour toute hétérogénéité observée dans les résultats de la revue, et une discussion sur celle-ci ?
15	S'ils ont réalisé une synthèse quantitative, les auteurs ont-ils mené une évaluation adéquate des biais de publication (biais de petite étude) et ont discuté de son impact probable sur les résultats de la revue ?
16	Les auteurs ont-ils rapporté toute source potentielle de conflit d'intérêts, y compris tout financement reçu pour réaliser la revue ?

Annexe 3 : L'échelle QUADAS-2

Phase 1 : Posez la question clinique :

Patients (contexte, motif de consultation, tableau clinique, tests précédemment réalisés) :

Test(s) évalué(s) :

Test de référence et condition cible :

Phase 2 : Dessinez le diagramme de flux de l'étude primaire

Phase 3 : Risque de biais et jugements d'applicabilité

La QUADAS-2 est organisée en 4 domaines qui sont chacun évalués en termes de risque de biais et d'applicabilité clinique de la question de recherche (telle que définie plus haut). Chaque domaine comprend une série de questions qui permettent de porter un jugement sur le risque de biais et l'applicabilité clinique.

DOMAINE 1 : SELECTION DES PATIENTS

A. Risque de biais

Décrivez la méthode de sélection des patients :

- Un échantillon de patients a-t-il été recruté de manière consécutive ou aléatoire ? Oui/Non/Incertain
- Un schéma d'étude type cas-témoins a-t-il été évité ? Oui/Non/Incertain
- L'étude a-t-elle évité des exclusions inappropriées ? Oui/Non/Incertain

La sélection des patients a-t-elle pu introduire un biais ?
RISQUE : FAIBLE/ELEVÉ/INCERTAIN

B. Préoccupations sur l'applicabilité

Décrivez les patients inclus (test préalable, présentation, utilisation prévue du test évalué, cadre) :

Y a-t-il un risque que les patients sélectionnés ne correspondent pas à la question de recherche ?
PREOCCUPATION : FAIBLE/ELEVEE/INCERTAINE

DOMAINE 2 : TEST EVALUE
Si plus d'un test est évalué, veuillez répondre à ces questions pour chaque test

A. Risque de biais

Décrivez le test évalué, comment il a été réalisé et interprété :

- Les résultats du test évalué ont-il été interprétés sans avoir connaissance des résultats du test de référence ? Oui/Non/Incertain
- Si un seuil de décision a été utilisé, a-t-il été spécifié ? Oui/Non/Incertain

La réalisation ou l'interprétation du test évalué a-t-elle pu introduire un biais ?
RISQUE : FAIBLE/ELEVÉ/INCERTAIN

B. Préoccupations sur l'applicabilité

Y a-t-il une préoccupation quant au fait que le test évalué, dans sa réalisation, interprétation, diffère de la question de recherche ?
PREOCCUPATION : FAIBLE/ELEVEE/INCERTAINE

DOMAINE 3 : TEST DE REFERENCE

A. Risque de biais

Décrivez le test de référence, comment il a été réalisé et interprété :

- Le test de référence est-il susceptible de correctement classer la condition recherchée ? Oui/Non/Incertain
- Les résultats du test de référence ont-ils été interprétés sans avoir connaissance des résultats du test évalué ? Oui/Non/Incertain

La réalisation ou l'interprétation du test de référence a-t-elle pu introduire un biais ?
RISQUE : FAIBLE/ELEVÉ/INCERTAIN

B. Préoccupations sur l'applicabilité

Y a-t-il une préoccupation quant au fait que la pathologie telle qu'elle a été diagnostiquée par le test de référence diffère de celle de la question de recherche ?
PREOCCUPATION : FAIBLE/ELEVEE/INCERTAINE

DOMAINE 4 : DEROULEMENT ET TEMPORALITE

A. Risque de biais

Décrivez tous les patients qui n'ont pas reçu le test évalué et/ou le test de référence ou qui ont été exclus du tableau de contingence 2X2 (confère au diagramme de flux) :

Décrivez l'intervalle de temps utilisé et tout traitement mis en place entre le test à évaluer et le test de référence :

- L'intervalle de temps entre la réalisation du test évalué et le test de référence est-il approprié ? Oui/Non/Incertain
- Tous les patients ont-ils reçu le test de référence ? Oui/Non/Incertain
- Tous les patients ont-ils reçu le même test de référence ? Oui/Non/Incertain
- Tous les patients sont-ils inclus dans l'analyse ? Oui/Non/Incertain

Le déroulement des tests auprès des patients a-t-il pu introduire un biais ?
RISQUE : FAIBLE/ELEVÉ/INCERTAIN

Annexe 4 : L'échelle QAREL

Tableau I. Traduction française de QAREL.

Items	Cotation
1	Le test a-t-il été évalué avec un échantillon de sujets représentatifs de la population à laquelle les auteurs ont l'intention d'appliquer les résultats ? Oui/Non/Incertain
2	Le test a-t-il été évalué par des évaluateurs représentatifs de ceux à qui les auteurs ont l'intention d'appliquer les résultats ? Oui/Non/Incertain
3	Les évaluateurs étaient-ils en aveugle par rapport aux résultats des autres évaluateurs au cours de l'étude ? Oui/Non/Incertain/Non applicable
4	Les évaluateurs étaient-ils en aveugle de leurs propres résultats antérieurs au test en cours d'évaluation ? Oui/Non/Incertain/Non applicable
5	Les évaluateurs étaient-ils en aveugle du résultat de l'étalon-or ou du statut de la pathologie pour la trouble (ou variable) évalué ? Oui/Non/Incertain/Non applicable
6	Les évaluateurs étaient-ils en aveugle par rapport à des informations cliniques non destinées à faire partie de la procédure de test ou du protocole de l'étude ? Oui/Non/Incertain/Non applicable
7	Les évaluateurs étaient-ils en aveugle par rapport à d'éventuels autres indices qui ne faisaient pas partie du test ? Oui/Non/Incertain
8	L'ordre d'évaluation variait-il ? Oui/Non/Incertain/Non applicable
9	Est-ce que la stabilité (ou la stabilité théorique) de la variable mesurée a été prise en compte pour déterminer la pertinence de l'intervalle de temps entre les mesures répétées ? Oui/Non/Incertain
10	Le test a-t-il été appliqué correctement et interprété de manière appropriée ? Oui/Non/Incertain
11	Est-ce que des mesures statistiques appropriées de l'accord entre les évaluateurs ont été utilisées ? Oui/Non/Incertain

Résumé

Contexte : Il existe peu d'outils permettant le diagnostic du Trouble Développemental de la Coordination valide et fiable. Le TDC est ainsi sous-évalué. Le DCDQ et le MABC-2 sont beaucoup étudiés. Ils sont différents par leur structure, leur administrateur et l'aspect du trouble qui est jugé. Le kinésithérapeute est absent du parcours de santé de ces troubles moteurs. **Objectifs :** Le but de cette revue systématique est de déterminer la validité et la fiabilité du DCDQ et du MABC-2 et faire un état des lieux du pouvoir diagnostic des professionnels dans la littérature scientifique ainsi qu'examiner son amélioration possible. **Méthode :** Une recherche bibliographique a été menée de Septembre 2019 à Mars 2020 sur différentes bases de données : PubMed, PEDro, Cochrane Library et Science Direct grâce à une équation de recherche spécifique élaborée par le modèle PICO suivant des critères d'éligibilité prédéfinis pour sélectionner des articles. La sélection des études suit des critères d'inclusions et d'exclusions prédéfinis. **Résultats :** Le résultat des 22 études montrent une bonne validité dans une population à risque pour les deux outils. Ils sont significativement corrélés avec le MABC et le BOT-2. La validité du DCDQ est correcte et celle du MABC-2 est modérée mais reste insuffisante dans la population globale. Les résultats du MABC-2 montrent une excellente cohérence interne ainsi qu'une bonne fiabilité statistique test-retest et inter-évaluateurs. Les résultats du DCDQ montrent lui aussi une excellente cohérence interne ainsi qu'une bonne fiabilité test-retest mais manque d'études concernant la fiabilité inter-évaluateurs. La fiabilité est assez bonne mais encore pas assez étudiée. **Conclusions :** Les deux outils ne sont pas assez efficaces pour une utilisation seule et un diagnostic optimal mais ils demeurent acceptés grâce à leur ratio qualité statistique/facilité d'administration/coût. Leur association peut être une option mais reste à étudier. Le diagnostic doit être ouvert aux kinésithérapeutes et l'administration du DCDQ par un professionnel doit être également étudiée.

Mots-Clés : DCDQ, MABC-2, Troubles moteurs, Performance motrice, Développement de l'enfant, Troubles développementaux de la Coordination, Validité, Fiabilité, Sensibilité, Spécificité, Diagnostic

Abstract

Context : There are few tools available for the diagnosis of Developmental Coordination Disorder that are valid and reliable. As a result, DCD is under-evaluated. DCDQ and MABC-2 are widely studied. They are different in their structure, their administrator and the aspect of the disorder that is judged. The physiotherapist is absent from the health course of these motor disorders. **Objectives :** The purpose of this systematic review is to determine the validity and reliability of DCDQ and MABC-2 and to take stock of the diagnostic power of professionals in the scientific literature as well as to examine its possible improvement. **Method :** A literature search was conducted from September 2019 to March 2020 on different databases: PubMed, PEDro, Cochrane Library and Science Direct using a specific search

equation developed by the PICO model according to predefined eligibility criteria to select articles. The selection of studies follows predefined inclusion and exclusion criteria. **Results :** The results of the 22 studies show good validity in an at-risk population for both tools. They are significantly correlated with MABC and BOT-2. The validity of the DCDQ is correct and that of the MABC-2 is moderate but remains insufficient in the overall population. The results of the MABC-2 show excellent internal consistency as well as good test-retest and inter-rater reliability. The DCDQ results also show excellent internal consistency and good test-retest reliability but lack studies on inter-rater reliability. Reliability is quite good but still under-researched. **Conclusions :** The two tools are not effective enough for use alone and optimal diagnosis but they remain accepted because of their statistical quality/easiness of administration/cost ratio. Their combination may be an option but remains to be studied. The diagnosis should be open to physiotherapists and the administration of the DCDQ by a professional should also be studied.

Key Words : DCDQ, MABC-2, Motor Skills Disorders, Motor Performance, Child Development, Developmental Coordination Disorders, Validity, Reliability, Sensibility, Specificity, Diagnosis