



HAL
open science

Fiabilité des outils d'évaluation de la proprioception chez les patients victimes d'un accident vasculaire cérébral

Adrien Penisson

► **To cite this version:**

Adrien Penisson. Fiabilité des outils d'évaluation de la proprioception chez les patients victimes d'un accident vasculaire cérébral. Médecine humaine et pathologie. 2021. dumas-03358211

HAL Id: dumas-03358211

<https://dumas.ccsd.cnrs.fr/dumas-03358211>

Submitted on 29 Sep 2021

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**

**FIABILITE DES OUTILS D'ÉVALUATION DE
LA PROPRICEPTION CHEZ LES PATIENTS
VICTIMES D'UN ACCIDENT VASCULAIRE
CEREBRAL**

Étudiant : PENISSON Adrien

Directeur de mémoire : ROSTAGNO Stephan

REMERCIEMENTS

Je remercie mon directeur de mémoire, Stephan Rostagno, qui a toujours su trouver le temps de me conseiller et de répondre à mes questions durant la rédaction de ce mémoire lors de nos nombreux entretiens.

Je remercie Cyril Duclos qui m'a proposé cette opportunité de publication dont peu d'étudiants ont la chance de profiter.

Je remercie Karim Jamal, avec qui j'ai passé de nombreuses heures à travailler sur ce mémoire et l'article assorti. Tu as souvent réussi me rassurer et me guider. J'espère sincèrement que nous aurons l'occasion de nous rencontrer un jour.

Je remercie Angèle, qui me supporte depuis quelques temps déjà. Tu as toujours su m'écouter râler et désespérer à propos de ce mémoire, et de tant d'autres choses. J'ai hâte d'ouvrir un nouveau chapitre de ma vie à tes côtés.

Enfin, je remercie ceux qui ont toujours été là et qui, je le sais, le seront jusqu'à la fin des temps : maman et papa. Aujourd'hui une aventure s'achève et une autre va bientôt commencer. Grâce à vous, je peux désormais voler de mes propres ailes. Je suis heureux d'avoir pu vous rendre fiers. Merci. Pour tout.

LISTE DES ABREVIATIONS

AMSTAR	<i>Assessing the Methodological Quality of Systematic Reviews</i>
AVC	Accident Vasculaire Cérébral
EmNSA	<i>Erasmus MC Modifications to the Nottingham Sensory Assessment</i>
FMA	<i>Fugl-Meyer Assessment</i>
FNM	Fuseau Neuro Musculaire
GRADE	<i>Grading of Recommendations Assessment, Development and Evaluation</i>
ICC	Coefficient de Corrélation Intraclasse
JPS	<i>Joint Position Sense</i>
KINARM	<i>Kinesiological Instrument for Normal and Altered Reaching Movements</i>
LEPT	<i>Lower Extremity Position Test</i>
MDC	<i>Minimal Detectable Change</i>
MeSH	<i>Medical Subject Headings</i>
NSA	<i>Nottingham Sensory Assessment</i>
OTG	Organe Tendineux de Golgi
QAREL	<i>Quality Appraisal of Diagnostic Reliability</i>
QUADRAS	<i>Quality Assessment of Diagnostic Accuracy Studies</i>
RASP	<i>Rivermead Assessment of Somato-sensory Perception</i>
rNSA	<i>Revised Nottingham Sensory Assessment</i>
SEM	<i>Standard error of Measurement</i>
TDM	<i>Threshold for Detection Motion</i>
TLT	<i>Thumb Localizing Test</i>
WPST	<i>Wrist Position Sense Test</i>

LISTE DES FIGURES

Figure 1 - Organisation d'un fuseau neuromusculaire.....	3
Figure 2 - Réponse de fibres afférentes primaires et secondaires à l'étirement musculaire	4
Figure 3 - Organisation d'un organe tendineux de Golgi	5
Figure 4 - La voie lemniscale	5
Figure 5 - KINARM Exoskeleton Lab.....	9
Figure 6 - Dynamomètre isocinétique.....	9
Figure 7 - Table magnétique à acquisition informatisée.....	9
Figure 8 - Organisation simplifiée du système vasculaire de l'encéphale	10
Figure 9 - Taxonomie des indices clinimétriques	14
Figure 10 - Recherche PubMed	17
Figure 11 - Recherche Cochrane	18
Figure 12 - Recherche ClinicalTrials.....	18
Figure 13 - Recherche Kinedoc.....	19
Figure 14 - Diagramme de flux	22

LISTE DES TABLEAUX

Tableau 1 - Interprétation du Coefficient de Corrélacion Intraclasse selon Koo	15
Tableau 2 - Interprétation du Coefficient de Kappa selon Landis et Koch.....	15
Tableau 3 - Interprétation de l'Alpha de Cronbach selon Nunnally.....	16
Tableau 4 - Caractéristiques des études incluses dans la revue	24
Tableau 5 - Récapitulatif de l'évaluation de la qualité méthodologique des articles par l'échelle QAREL	26
Tableau 6 - Résultats des études sur la fiabilité inter-évaluateur	31
Tableau 7 - Résultats des études sur la fiabilité intra-évaluateur	34
Tableau 8 - Résultats des études sur l'Erreur de Mesure Standard.....	35
Tableau 9 - Synthèse des résultats obtenus par la revue et leurs niveaux de preuve analysés par le système GRADE	47
Tableau 10 - Évaluation de la qualité méthodologique de la présente revue de littérature ..	50

TABLE DES MATIERES

1 – INTRODUCTION	1
1.1 La Proprioception	1
1.1.1 Généralités	1
1.1.2 Les récepteurs périphériques proprioceptifs.....	2
1.1.3 La centralisation de la proprioception	5
1.1.4 Les méthodes d'évaluation de la proprioception post-AVC	6
1.2 L'Accident Vasculaire Cérébral	10
1.2.1 Physiopathologie.....	10
1.2.2 Rééducation du patient victime d'un AVC	12
1.3 Objectif de la revue de littérature	12
2 – METHODE DE RECHERCHE.....	13
2.1 Critères d'éligibilité des études	13
2.1.1 Schéma d'étude	13
2.1.2 Population et pathologie étudiées.....	13
2.1.3 Outils de mesure évalués.....	13
2.1.4 Critères de jugement	13
2.2 Méthode de recherche des études	16
2.2.1 Sources documentaires investiguées.....	16
2.2.2 Équation de recherche utilisée	17
2.2.3 Évaluation de la qualité méthodologique des études sélectionnées	19
2.3 Méthode d'extraction des données et de synthèse des résultats.....	19
2.3.1 Méthode de sélection des études.....	19
2.3.2 Méthode d'extraction des données.....	20
3 – RESULTATS	22
3.1 Sélection des études	22
3.1.1 Diagramme de flux.....	22
3.1.2 Études exclues.....	23
3.1.3 Études retenues	23
3.2 Extraction des données	23
3.3 Analyse des données des études retenues	25
3.3.1 Qualité méthodologique des articles.....	25
3.3.2 Caractéristiques générales de la population	26
3.3.3 Outils d'évaluation de la proprioception	27
3.3.4 Articulations évaluées.....	27
3.3.5 Propriétés métrologiques étudiées	27
3.3.6 Analyse descriptive des données.....	28
4 – DISCUSSION.....	36
4.1 Analyse et qualité des preuves des principaux résultats	36
4.1.1 La fiabilité inter-évaluateur.....	36
4.1.2 La fiabilité intra-évaluateur.....	40
4.1.3 La consistance interne et l'erreur de mesure standard.....	42
4.2 Niveaux de preuve	43
4.3 Applicabilité des résultats en pratique.....	48

4.4 Biais potentiels de la revue	49
5 – CONCLUSION	52
BIBLIOGRAPHIE.....	I
ANNEXES	VIII

1 – INTRODUCTION

On dénombre chaque année en France environ 140 000 cas d'Accidents Vasculaires Cérébraux (AVC). Avec des séquelles importantes dans 40% des cas, ils représentent la première cause de handicap moteur acquis de l'adulte ainsi que la deuxième cause de mortalité dans notre pays, et leur risque de récurrence est estimé à 10% dans les cinq années suivantes. C'est donc une pathologie fréquente et potentiellement grave constituant une urgence médicale.¹ [1]

L'AVC, nous le détaillerons plus loin, est souvent responsable de déficits moteurs et sensitifs, et par conséquent de troubles proprioceptifs, c'est-à-dire de la sensibilité dite profonde. Plus précisément, entre 34 et 64% des patients victimes d'un AVC présentent des troubles de la proprioception [2], or cette dernière est un facteur essentiel pour leur récupération motrice fonctionnelle [3]. Une bonne évaluation de la sévérité des troubles proprioceptifs est donc primordiale pour proposer une rééducation de qualité chez ces patients, cependant il existe de nombreux outils permettant d'effectuer un bilan proprioceptif [4], aussi bien manuels que techniques. Nous allons tenter par ce travail de voir s'il existe une différence significative de fiabilité entre certains de ces outils dans la prise en charge des patients.

1.1 La Proprioception

1.1.1 Généralités

Afin d'être en pleine mesure de planifier et d'effectuer des mouvements adaptés aux situations rencontrées, l'être humain doit avant tout être capable de percevoir son environnement et ses variations. Pour ce faire, 3 principaux systèmes sensoriels entrent en jeu [5] :

Le système visuel, dans lequel l'œil va capter des signaux lumineux pour les transformer en signaux électriques à destination du nerf optique puis du cerveau, qui nous renvoie lui-même les informations reçues sous la forme d'une image que nous interprétons. [6]

Le système vestibulaire, situé dans l'oreille interne et jouant un rôle important dans le contrôle de l'équilibre. [7]

Le système somatosensoriel [8], [9], composé lui-même de :

- La sensibilité tactile, nous informant sur la perception du toucher et de toutes ses variations
- La sensibilité nociceptive (la douleur) et thermique (les changements de température)
- La sensibilité proprioceptive, détaillée ci-dessous

¹ Inserm, "Accident vasculaire cérébral," 2019. [Online]. Available: <https://www.inserm.fr/information-en-sante/dossiers-information/accident-vasculaire-cerebral-avc>

La proprioception est une sensibilité à la fois consciente et inconsciente que certains qualifient même de 6^{ème} sens [10]. Dans ce travail, nous développerons la proprioception dite consciente, et nous nous attarderons plus particulièrement sur deux de ses aspects : la statestésie et la kinesthésie. [10], [11]

La proprioception est le sens de la position du corps et de ses membres (c'est la statestésie) et de leurs mouvements (kinesthésie). L'information proprioceptive provient des récepteurs périphériques, détaillés plus bas, et contribue au contrôle postural, à la stabilité des articulations, à la coordination et à l'appréciation de la position et du déplacement des membres dans l'espace. Elle participe aussi au maintien de l'équilibre, en parallèle du système visuel et du système vestibulaire. En dehors de l'équilibre, elle joue un rôle essentiel dans les activités motrices fines, comme l'écriture. [4]

Nous nous intéressons ici aux deux principaux types de proprioception : la conscience du statique, ou statestésie, et la conscience du mouvement ou kinesthésie. Ces deux types de sensibilité profonde dépendent de récepteurs périphériques communs. Les différences perceptibles entre la statestésie et la kinesthésie peuvent donc être attribuées aux récepteurs périphériques, qu'ils soient musculaires ou articulaires. [11]

1.1.2 Les récepteurs périphériques proprioceptifs

Pour parvenir à conserver un bon équilibre et à effectuer des mouvements, l'organisme utilise des systèmes de contrôle, d'une part extéroceptifs et d'autre part proprioceptifs en utilisant différents récepteurs sensitifs. [12]

Les récepteurs sensitifs extéroceptifs traitent les informations extérieures au corps, telles que la vue, l'ouïe, l'odorat, le toucher et le goût. Les récepteurs proprioceptifs sont les récepteurs de la sensibilité profonde, ils sont internes et propres à notre corps ; ils siègent dans les muscles et les articulations. Selon le type de récepteurs proprioceptifs stimulés, la proprioception permet d'évaluer la position d'un segment de corps dans l'espace, soit de façon statique, soit de façon dynamique. [13]

1.1.2.1 Les Fuseaux Neuro Musculaires

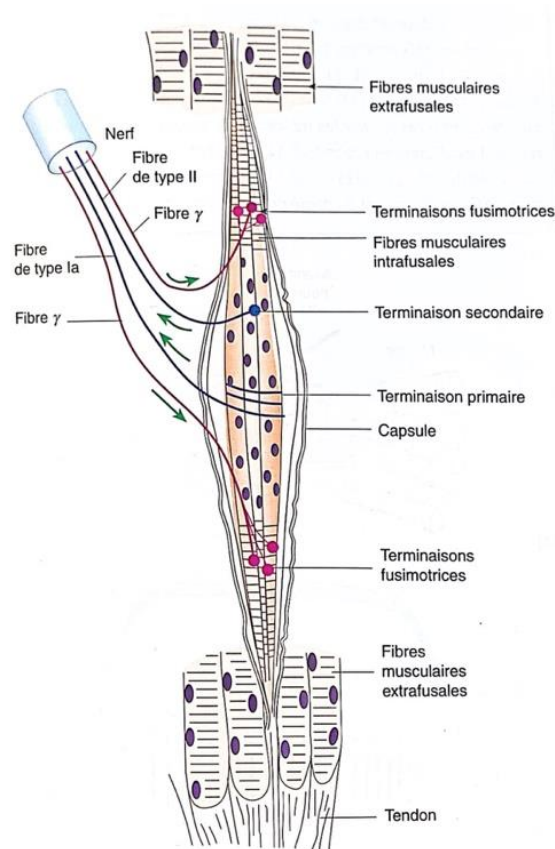


Figure 1 - Organisation d'un fuseau neuromusculaire

Les FNM (voir figure 1) sont considérés comme les principaux récepteurs proprioceptifs et sont notamment retrouvés en grande quantité dans les muscles de la motricité fine. Chaque FNM se compose de faisceaux, eux-mêmes composés de fibres musculaires enfermées dans une capsule de tissu conjonctif et innervées par des fibres motrices efférentes et des fibres sensibles afférentes. Dans un muscle squelettique, on va donc retrouver deux types de fibres musculaires, les fibres du fuseau à proprement parlé, elles sont dites intrafusales, et les fibres musculaires en dehors de la capsule conjonctive, donc dites extrafusales.

Les fibres intrafusales se divisent elles-mêmes en deux catégories, d'une part les fibres à sacs nucléaires et d'autre part les fibres à chaînes nucléaires. La principale différence est que les premières comportent des groupements de noyaux nucléaires, alors que les secondes présentent un agencement en série de leurs noyaux et sont un peu plus petites.

Au niveau des efférences, les motoneurones innervant les fibres intrafusales sont appelés motoneurones gamma, afin de les différencier des motoneurones alpha, innervant les fibres musculaires extrafusales. Les fibres des motoneurones gamma, appelées fibres fusimotrices, ont un diamètre et une vitesse de conduction bien moindre que les fibres des motoneurones alpha, bien qu'elles soient également myélinisées, et proviennent de la corne ventrale de la moelle épinière.

Enfin, l'afférence des FNM est assurée par deux types de fibres, qui permettent là encore de diviser les FNM en deux catégories, primaires (afférences Ia) et secondaires (afférences II). Les fibres afférentes primaires, myélinisées et de gros calibres, et forment des terminaisons dites annulo-spiralées, et se retrouvent plutôt sur la partie médiane des fibres intrafusales à sacs nucléaires et à chaînes nucléaires. Les fibres afférentes secondaires, elles, forment des terminaisons en bouquet, c'est-à-dire à ramifications multiples, et se retrouvent plus en périphérie des fibres intrafusales, et en majorité sur les fibres à chaînes nucléaires.

Les afférences de type Ia sont beaucoup plus sensibles à la tension appliquée sur la fibre musculaire que les afférences de type II (voir figure 2), c'est pour cela qu'on considère que les primaires sont dynamiques et que les secondaires sont statiques. Les FNM primaires vont indiquer la vitesse de changement d'état du muscle (contraction, relâchement, étirement) et vont donc nous informer sur le déroulement d'un mouvement effectué (ils participent aussi au réflexe myotatique), alors que les FNM secondaires donnent des informations sur la longueur du muscle et donc sur la position d'un segment de membre.

[13]–[15]

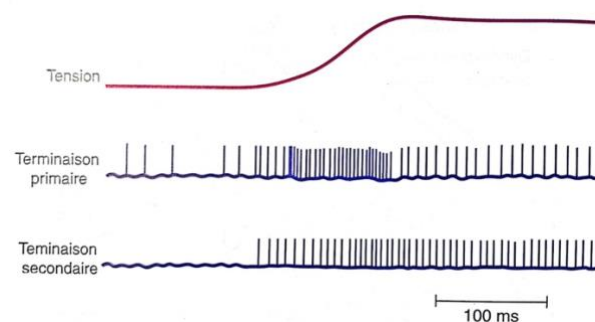


Figure 2 - Réponse de fibres afférentes primaires et secondaires à l'étirement musculaire

1.1.2.2 Les Organes Tendineux de Golgi

Les organes tendineux de Golgi (voir figure 3) siègent au niveau des jonctions tendineuses, ils sont situés en série par rapport aux fibres extrafusales évoquées plus haut. Ils sont sensibles à la tension intra musculaire et participent à la régulation de la raideur musculaire et au réflexe myotatique inverse. Ils se situent également au niveau des ligaments où ils signalent la tension appliquée à ces derniers et la direction du mouvement. Ils ont aussi un rôle statique, en participant à la statesthésie [15]. Ils sont également retrouvés dans les ligaments. Leur afférence est assurée par des neurones de type Ib, myélinisés et d'un calibre similaire aux fibres Ia. Les fibres afférentes Ib pénètrent dans la moelle épinière au niveau des racines dorsales et forment des synapses avec des interneurones au niveau de la corne ventrale, qui vont à leur tour former des synapses avec des motoneurones alpha ; les connexions ainsi créées provoquent un relâchement musculaire lorsqu'elles sont activées [16]. A l'opposé des FNM, les OTG n'ont pas d'afférences directes avec le système nerveux central, et ne sont donc pas sujets à ses atteintes.

La fréquence des potentiels d'action dans l'afférence Ib est surtout augmentée par la force de contraction du muscle, alors l'étirement passif de ce dernier ne produit qu'une faible réponse des OTG. [15]

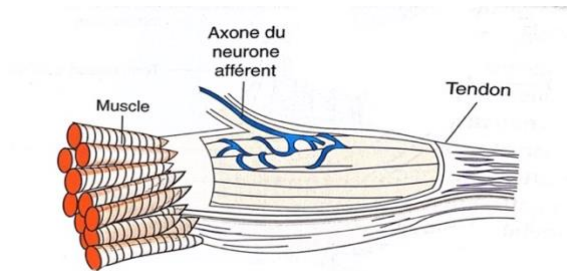


Figure 3 - Organisation d'un organe tendineux de Golgi

Il existe d'autres récepteurs proprioceptifs moins importants, comme les récepteurs de Ruffini, retrouvés dans les capsules articulaires, ou encore les corpusculaires de Pacini [17].

1.1.3 La centralisation de la proprioception

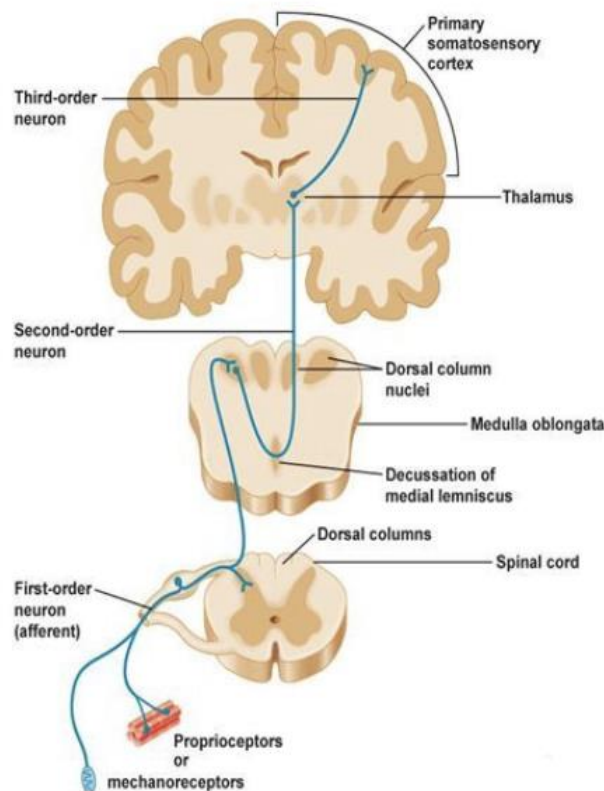


Figure 4 - La voie lemniscale

La proprioception est liée aux voies somatosensorielles, et plus précisément à la voie lemniscale (voir figure 4). En effet une fois les récepteurs proprioceptifs stimulés, les fibres nerveuses afférentes, primaires ou secondaires, vont faire circuler l'information en pénétrant dans la moelle épinière par les racines postérieures. Les afférences dites de premier ordre se déplaceront ainsi par les colonnes dorsales pour former des synapses avec les noyaux du cordon postérieur (noyaux cunéiforme et gracile) du bulbe rachidien. Dans un second temps, les afférences de deuxième ordre vont quitter ces mêmes noyaux dans un faisceau de fibres appelé lémnisque médial avant d'atteindre le thalamus ventral. Enfin, les fibres de troisième ordre, issues de ce dernier, vont rejoindre les régions somesthésiques du cortex cérébral. Le cortex somatosensoriel primaire droit va traiter l'information provenant de l'hémicorps gauche, et inversement. [15], [18]

1.1.4 Les méthodes d'évaluation de la proprioception post-AVC

1.1.4.1 Généralités

Comme expliqué précédemment, jusqu'à 2/3 des patients présentent des troubles proprioceptifs à la suite d'un AVC, et la sensibilité proprioceptive est indispensable à une bonne récupération motrice et de l'équilibre. Sa prise en charge étant une étape clé dans la réhabilitation du patient, il est fondamental de pouvoir l'évaluer de manière précise et objective, afin de pouvoir identifier les déficits induits et proposer un plan de traitement adapté au malade dans sa prise en charge. [1], [2]

Les kinésithérapeutes semblent accorder de plus en plus d'importance à l'évaluation de la proprioception chez les patients post-AVC, en effet 95% des kinésithérapeutes interrogés dans une étude de 2015 [19] affirment évaluer les déficits proprioceptifs de leurs patients post-AVC, contre 84% dans une étude de 1999 [20]. Bien que les deux études n'aient pas été menées auprès des mêmes groupes de thérapeutes, la plus récente émet l'hypothèse que cette augmentation de la préoccupation des rééducateurs envers les troubles somatosensoriels est due à la large diffusion des données scientifiques probantes relatives au sujet et au développement des guides de bonnes pratiques.

A ce jour, il n'existe pas de réel Gold Standard (GS) pour l'évaluation de la proprioception, même si certaines échelles sont plus utilisées que d'autres en pratique clinique. Certaines méthodes, souvent robotisées mais pas toujours, sont parfois considérées comme étant des GS par certains auteurs dans le cas d'études comparatives. [21]

1.1.4.2 Les méthodes manuelles

Le **Joint Position Sense (JPS)**, pour la stathésie [22], [23] : il permet de mesurer en passif ou en actif la capacité à détecter la position d'un membre dans l'espace. Cet examen consiste à placer l'articulation passivement dans une position et on demande au sujet de placer la même articulation controlatérale dans la même position, les yeux fermés. Cet

examen a été décliné en d'autres plus spécifiques comme le **Thumb Localizing Test** [24] ou encore le **Wrist Position Sense Test** [25].

Le **Threshold for Detection Motion** (TDM), pour la kinesthésie [22] : il consiste à mesurer la capacité à ressentir le mouvement à partir d'une phase immobile. Pour ce faire, on va imprimer, en passif et articulation par articulation une séquence de mouvements au sujet et, les yeux fermés, celui-ci doit effectuer cette séquence en actif avec le membre controlatéral.

Ces 2 tests sont la base de l'évaluation de la proprioception et leurs principes sont retrouvés dans des examens de la sensibilité plus vastes. Voici les plus utilisés :

Le **Nottingham Sensory Assessment** (NSA) est une évaluation manuelle standardisée utilisée dans de nombreux pays pour les patients présentant des troubles sensitifs, et notamment en post-AVC. C'est une échelle britannique datant de 1991 et incluant les items suivants : toucher léger, température, piqûre, pression, localisation tactile, stimulation bilatérale cutanée, stéréognosie, proprioception, discrimination de deux points [26]. Sept ans plus tard, une version révisée et plus courte du NSA a vu le jour, le **Revised Nottingham Sensory Assessment** (rNSA) [27]. Enfin, neuf ans plus tard, en 2007, une version actualisée du rNSA a été publiée : le **Erasmus MC Modifications to the Nottingham Sensory Assessment** (EmNSA) [28]. Si toutes les modifications effectuées au fil des versions de la NSA ne seront pas détaillées dans ce travail, attardons-nous tout de même sur celles qui concernent la partie proprioception (pour rappel, chaque test doit être effectué 3 fois yeux fermés pour chaque articulation visée).

- Dans le NSA la partie proprioception se décomposait en 3 catégories à évaluer : l'appréciation du mouvement, la direction du mouvement et la position de l'articulation. Les scores possibles étaient les suivants :
 - 0 Absente : Incorrect lors des 3 essais
 - 1 Atteinte : Correct lors de 1 ou 2 essai(s)
 - 2 Normale : Correct lors des 3 essais
- Dans le rNSA la partie proprioception a été totalement remaniée, en effet les 3 catégories évoquées précédemment disparaissent pour devenir à la place des critères de notation :
 - 0 Absente : Aucune appréciation du mouvement
 - 1 Appréciation du mouvement : Le patient indique lors des 3 essais qu'il y a un mouvement mais la direction est incorrecte
 - 2 Sensation de direction du mouvement : Le patient est capable d'apprécier et mimer la direction du mouvement à chaque fois, mais la position est incorrecte
 - 3 Sensation de position de l'articulation : Le patient parvient à reproduire le mouvement à 10 degrés près de la position du test
- Enfin, dans le EmNSA, la partie proprioception a été simplifiée en supprimant toute information relative à la sensation de position de l'articulation :
 - 0 Absente : Le patient ne détecte aucun mouvement

- 1 Atteinte : Le patient détecte le mouvement mais la direction n'est pas correcte lors des 3 essais
- 2 Normale : le patient détecte correctement la direction du mouvement lors des 3 essais

Le **Fugl-Meyer Assessment (FMA)** est une échelle d'évaluation à la fois sensitive et motrice créée pour les patients présentant des troubles hémiparétiques suite à un AVC développée en 1975 [29]. Il comporte 5 catégories (fonction motrice, sensibilité, équilibre, amplitude articulaire et douleur articulaire) et 155 items au total. Dans la sous-catégorie traitant de la proprioception, chaque articulation est testée 4 fois en comparaison avec le côté sain, les yeux fermés.

- 0 : moins de 3/4 correct
- 1 : 3/4 correct
- 2 : 4/4 correct

En 2007, une version courte de 12 items a vu le jour, mais centrée uniquement sur la partie motrice [30]. Le FMA est évoqué dans le bilan sensitif du patient hémiparétique par la Haute Autorité de Santé française².

Le **Rivermead Assessment of Somatosensory Performance (RASP)** est une échelle standardisée qui évalue la sensibilité d'après 7 catégories [31]. La dernière, dédiée à la proprioception, est divisée en deux : la détection de mouvement (réponses possibles pour le patient oui/non) et la détection de direction (haut/bas). Le test s'effectue les yeux fermés. La notation est la suivante :

- 0 : mauvaise réponse
- 1 : bonne réponse

Le RASP teste uniquement le coude, le poignet, les doigts, la cheville et les orteils, avec 6 essais chacun. Chaque essai est noté pour obtenir finalement un score total sur 30 par hémicorps. Si le patient obtient un score inférieur 28 les auteurs considèrent qu'il présente des troubles proprioceptifs.

1.1.4.3 Les méthodes instrumentales

Le **Kinesiological Instrument for Normal and Altered Reaching Movements (KINARM)** Exoskeleton Lab est un fauteuil roulant avec un exosquelette sur sa partie supérieure attaché au sujet (voir figure 5). C'est un outil robotique permettant d'évaluer la proprioception de l'épaule et du coude d'un patient dans de multiples pathologies, notamment en post-AVC [32]. Un des arguments commerciaux utilisés par cette machine est de proposer une évaluation proprioceptive plus fiable et plus précise que les autres méthodes, supprimant ainsi la

² HAS, « Évaluation fonctionnelle de l'AVC » in *Référentiel d'auto-évaluation des pratiques professionnelles en massokinésithérapie*, Janvier 2006. [Online]. Available : https://www.has-sante.fr/upload/docs/application/pdf/Evaluation_%20fonctionnelle_%20AVC_rf.pdf



Figure 5 - KINARM Exoskeleton Lab

subjectivité inhérente au thérapeute qui effectue le bilan. Couplé à la réalité augmentée, le KINARM peut également permettre d'évaluer et quantifier les interactions entre les systèmes moteur, visuel et proprioceptif lors de l'exécution de diverses tâches.

Il s'agit d'une méthode très coûteuse, en effet après un échange par mail avec Anne Vivian-Scott, la présidente de la branche KINARM de BKIN Technology, le prix de l'instrument a été communiqué : 250 000 dollars canadiens à l'achat, avec un prix d'utilisation de la licence logiciel officielle de 6 500 dollars canadiens annuels (voir Annexe I). C'est un outil principalement destiné à la recherche, et non au diagnostic clinique, du fait de son coût et sa mise en place particulière.

D'autres méthodes instrumentales existent cependant, comme par exemple la mesure de la proprioception par **dynamomètre isocinétique** (voir figure 6) [33], qui consiste à mesurer le seuil de détection d'un mouvement grâce à la vitesse angulaire de l'instrument et l'erreur de positionnement de l'articulation testée par rapport à l'articulation controlatérale grâce à l'angle articulaire. Là encore, nous nous heurtons à un problème financier, puisque les entreprises développant ces machines les vendent à des sommes situées entre 60 000 et 150 000 euros, auxquelles il faut le plus souvent rajouter le prix d'utilisation d'une licence annuelle.



Figure 6 - Dynamomètre isocinétique

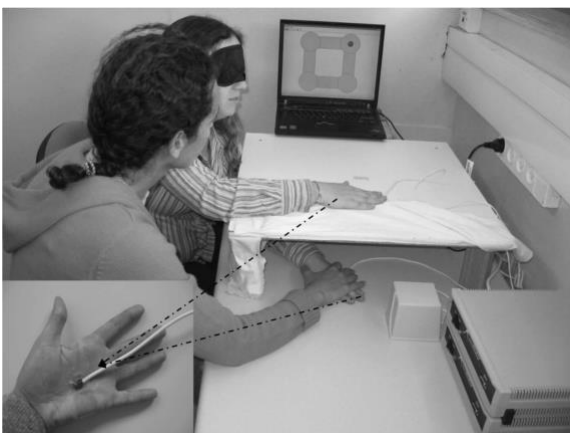


Figure 7 - Table magnétique à acquisition informatisée

Nous pouvons également évoquer ici une autre méthode de mesure informatisée, une **table magnétique** associée à des capteurs, développée il y a quelques années (voir figure 7) [34]. Ici, le patient doit reproduire sur une table magnétique des mouvements et/ou positions avec son membre atteint en conservant dans sa main un capteur responsable de l'acquisition des données. Aucune information quant au prix de cet instrument n'a été trouvée.

Les outils évoqués ici ne sont donnés qu'à titre indicatif car il s'agit des plus répandus, il existe certainement d'autres méthodes de mesures instrumentales, commercialisées ou encore en développement.

1.2 L'Accident Vasculaire Cérébral

1.2.1 Physiopathologie

D'après le Collège des Enseignants de Neurologie, les AVC sont caractérisés par la survenue brutale d'un déficit neurologique focal, à l'opposé d'atteintes diffuses progressives. [35]

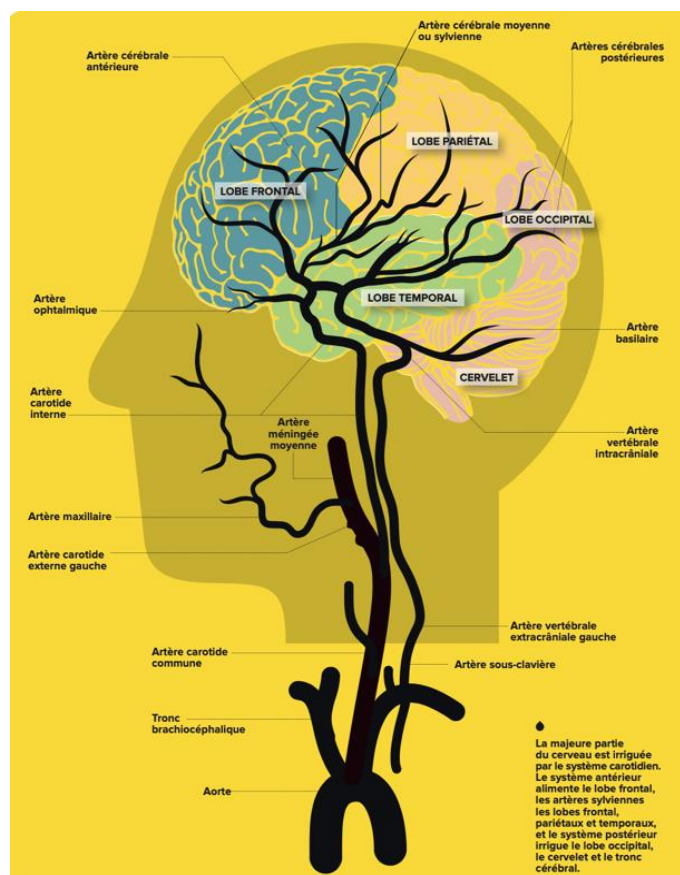


Figure 6 - Organisation simplifiée du système vasculaire de l'encéphale

Le terme d'AVC regroupe l'ensemble des pathologies vasculaires cérébrales d'origine artérielle ou veineuse qui vont provoquer une interruption d'oxygénation d'un territoire cérébral plus ou moins étendu, à savoir :

Les ischémies cérébrales artérielles (correspondant à environ 80% des AVC) ; généralement il s'agit d'une occlusion artérielle par thrombose ou embolie, mais, dans des cas plus rares, l'origine peut être hémodynamique, c'est-à-dire qu'on observe une chute de la

perfusion cérébrale sans occlusion ; c'est le cas lorsque la pression artérielle connaît une chute trop importante.

Les hémorragies cérébrales (correspondant à environ 20% des AVC), où une artère cérébrale se rompt et le sang se répand dans le parenchyme cérébral, y formant ainsi un hématome plus ou moins profond.

Les thromboses veineuses cérébrales (qui sont très rares, mais existantes).

[35]–[37]

En fonction du territoire cérébral atteint (voir figure 8), les déficiences varient et peuvent être cognitives, motrices et/ou sensorielles. Voici une liste, non exhaustive, des troubles fréquemment retrouvés chez un patient victime d'un AVC :

Parmi les atteintes cognitives principales, on retrouve les aphasies, qui sont des troubles du langage et/ou de la compréhension selon leur type. On peut également retrouver une démence chez certains patients. [38]

L'atteinte motrice la plus fréquente est l'hémi-parésie, voire l'hémiplégie [39], qui consiste en la perte partielle ou totale de la motricité d'un hémicorps sans atteinte nerveuse périphérique ou musculaire. [40]

Enfin, dans les atteintes sensorielles, on retrouve notamment l'hémianesthésie, qui constitue une perte de la sensibilité partielle ou totale d'un hémicorps, et nous nous intéressons principalement ici à l'anesthésie proprioceptive, qui est la perte du sens de position du corps dans l'espace. La voie lemniscale, responsable entre autres de la proprioception et présentée dans la *partie 1.1.3*, fait intervenir diverses structures et notamment, au niveau du cerveau, le cortex somesthésique primaire, situé dans le lobe pariétal, et le thalamus. Lorsqu'un AVC va priver d'oxygène une des structures précédemment évoquées, la voie lemniscale sera donc touchée et la chaîne d'informations rompue ou, à minima, abimée : le patient éprouvera des difficultés, voire une incapacité à situer des segments de son corps dans l'espace. C'est notamment le cas si l'artère cérébrale moyenne est touchée : le lobe pariétal, et donc le cortex somesthésique, ne sera plus irrigué. Parmi les atteintes touchant le thalamus, on retrouve le syndrome de Dejerine-Roussy qui touche entre 4 et 48% des patients victimes d'un AVC [41] et qui provoque des troubles sensitifs à type de dysesthésie ou allodynie, venant donc interférer dans la voie spino-thalamo-corticale, et ainsi dans le traitement de l'information proprioceptive.

Tous ces déficits provoqués par un AVC vont engendrer des séquelles sensorimotrices et/ou cognitives, qui vont impacter les activités de la vie quotidienne des patients, leur autonomie et leur réinsertion sociale. [42]

De nombreux facteurs de risques extérieurs sont évitables pour diminuer le risque d'AVC : l'obésité [43], [44], le cholestérol [45], le tabagisme [46], [47], la consommation d'alcool [48], etc. D'autres, plus intrinsèques aux individus, nécessitent une surveillance accrue, c'est le cas par exemple de l'hypertension artérielle [49] et de la fibrillation auriculaire [50].

En 2010, on déplorait 17 millions d'AVC dans l'année dans le monde, avec un âge moyen de 74 ans, mais dont plus d'un tiers touchait des personnes âgées de moins de 65 ans. Il a été estimé que le nombre annuel d'AVC dans le monde devrait passer le cap des 23 millions en 2030 du fait du vieillissement de la population [51]. D'ici 2025, le nombre de patients ayant survécu à un AVC est estimé à environ 77 millions à travers le monde, or comme dit plus haut, entre un et deux tiers d'entre eux présenteront des troubles de la proprioception. Cette dernière étant nécessaire à une bonne récupération motrice [52], son évaluation est un élément majeur de la bonne prise en charge des patients.

1.2.2 Rééducation du patient victime d'un AVC

La rééducation de l'AVC se décline en trois phases : la première est la phase aiguë parfois dite « flasque » qui s'étale sur les premiers jours/semaines après l'accident, et pendant laquelle le thérapeute lutte prioritairement contre les complications de l'immobilité et va travailler sur l'initiation de la commande motrice. S'en suit une phase de récupération, généralement de plusieurs mois, durant laquelle la prise en charge doit être intensive, afin de permettre une diminution des limitations d'activité du patient. On va maximiser le travail moteur et cognitif pour permettre au patient d'améliorer son autonomie, c'est un travail pluridisciplinaire. Enfin, le patient entre dans la phase de stabilisation, ou phase « séquellaire », durant laquelle on va axer la rééducation sur des objectifs plutôt précis dépendant des activités de la vie quotidienne de chaque patient pour augmenter leur qualité de vie ; bien sûr, il faut également conserver les progrès obtenus lors de la phase précédente. [53]–[55]

Et la proprioception dans tout ça ? Elle est généralement mise de côté dans la rééducation, les thérapeutes privilégiant la rééducation motrice pure. Il a cependant été prouvé que la proprioception entretenait un lien étroit avec les capacités motrices, que ce soit dans le contrôle de la marche, de l'équilibre ou dans les activités plus fines [56]. Il existe tout de même des exercices et méthodes à la disposition des rééducateurs comme les entraînements multi-sensorimoteurs [57], c'est-à-dire combinant à la fois des exercices d'équilibre et des exercices sensitifs par exemple, la thérapie miroir [58] ou plus récemment la réalité virtuelle [59] et les méthodes robots assistées [60]–[61], qui semblent avoir un intérêt.

1.3 Objectif de la revue de littérature

L'objectif de ce travail est de faire un état des lieux des connaissances actuelles en matière d'évaluation de la proprioception chez les sujets victimes d'un AVC, puisque celle-ci présente un enjeu majeur dans leur récupération. Ce travail aura pour but de comparer en priorité la fiabilité de différents outils de mesure entre eux, afin de déterminer lesquels sont les plus pertinents à utiliser pour effectuer un bon diagnostic clinique.

2 – METHODE DE RECHERCHE

2.1 Critères d'éligibilité des études

Pour rappel, la question posée par ce travail concerne les différents outils d'évaluation de la proprioception et leur fiabilité chez les sujets d'atteints d'AVC.

2.1.1 Schéma d'étude

D'après le modèle PICO [62], détaillé plus bas, la question posée par ce travail est une question diagnostique. La fiabilité d'un outil de mesure diagnostic s'étudie généralement avec des études transversales et/ou longitudinales [63]. Pour rappel, une étude transversale porte sur une observation d'un phénomène et/ou d'un échantillon de population à un moment précis. A contrario, une étude longitudinale consiste en l'observation répétée pendant une période définie d'un phénomène et/ou d'un échantillon d'individus, permettant ainsi de pouvoir analyser l'évolution au fil du temps.

2.1.2 Population et pathologie étudiées

Dans le cadre de ce travail, la population présentée dans les études retenues doit répondre aux critères d'inclusion suivants : chaque protocole doit présenter des patients adultes, hommes ou femmes, atteints d'un AVC hémorragique ou ischémique, avec ou sans déficits apparents, et aucune limite d'âge maximale n'a été retenue. Les critères d'exclusion correspondent aux études portant sur la proprioception du tronc et aux études centrées sur les nouveau-nés victimes d'un AVC périnatal.

Si plusieurs pathologies sont traitées dans un article, seules les données en lien avec les critères d'inclusion seront extraites. Si, parmi les résultats, les données sur les patients AVC ne sont pas disponibles isolément, l'article sera exclu de la revue.

2.1.3 Outils de mesure évalués

Les outils de mesure qui seront évalués sont les techniques permettant d'évaluer la proprioception chez un patient dans le cadre d'un bilan post-AVC. Il s'agit plus précisément d'outils de mesure utilisés dans le cadre de diagnostics cliniques ou de recherche, et ils correspondent à une hétéroévaluation, c'est-à-dire une évaluation réalisée non pas par le patient, mais par une tierce personne. [64]

2.1.4 Critères de jugement

Les critères de jugement étudiés dans ce travail seront relatifs à la fiabilité des outils d'évaluation précédemment évoqués. Pour définir ces critères, nous nous sommes appuyés

sur les recommandations du COSMIN et de leur représentation du consensus sur la taxonomie, la terminologie et la définition des propriétés de mesures (voir figure 9). [65]



Figure 7 - Taxonomie des indices clinimétriques

2.1.4.1 Critère de jugement principal

Dans le cadre de ce travail, le critère que nous chercherons principalement à évaluer est la **fiabilité (reliability)**, c'est-à-dire le degré auquel une mesure est exempte d'erreur de mesure. La fiabilité juge ainsi de la concordance d'un outil de mesure [65], à ne pas confondre avec la corrélation. [66]

Il existe 2 principaux types de fiabilité : la *fiabilité intra-examineur (intra-rater reliability)* et la *fiabilité inter-examineur (inter-rater reliability)*. La première étudie si, lorsqu'un unique examineur effectue plusieurs fois une même mesure à un temps d'intervalle, les valeurs obtenues sont identiques et, si non, dans quelles proportions elles sont différentes. La seconde étudie si, lorsque plusieurs examineurs effectuent une même mesure, les valeurs obtenues sont identiques et, si non, dans quelles proportions elles sont différentes. [64]

Pour juger de la fiabilité d'un outil de mesure, il existe plusieurs coefficients dont 2 principaux : le Coefficient de Corrélation Intraclasse (ICC) et le Coefficient Kappa de Cohen (Kappa κ). L'ICC est généralement utilisé pour évaluer la fiabilité d'outils de mesure à variable quantitative, alors que le Kappa permet plutôt d'évaluer la fiabilité d'outils de mesure à variable qualitative [64]. Il est important que le kappa soit pondéré, c'est-à-dire que son calcul permette d'apprécier correctement la concordance d'un outil à variable qualitative ordonnée, sans cela le kappa ne permet d'apprécier la concordance que des outils à variable qualitative binaire. [66]

Dans ce travail, l'interprétation de l'ICC sera effectuée selon Koo [67] (voir tableau 1) et l'interprétation du Kappa de Cohen selon Landis et Koch [68] (voir tableau 2).

Valeur ICC	Niveau d'accord
< 0,50	Faible
0,50 - 0,75	Moyen
0,76 - 0,90	Bon
> 0,90	Excellent

Tableau 1 - Interprétation du Coefficient de Corrélation Intraclasse selon Koo

Valeur du Kappa	Niveau d'accord
0,00 - 0,20	Mauvais
0,21 - 0,40	Faible
0,41 - 0,60	Moyen
0,61 - 0,80	Bon
> 0,80	Excellent

Tableau 2 - Interprétation du Coefficient de Kappa selon Landis et Koch

2.1.4.2 Critères de jugement secondaires

La **consistance interne** (appréciée à l'aide du coefficient alpha de Cronbach α) peut aussi être utilisée pour apprécier la fiabilité d'un outil. La consistance interne analyse la corrélation entre les items d'une même échelle : elle indique si les items du test sont homogènes, similaires dans leur contenu [64]. La mesure du coefficient de consistance interne ne requiert qu'une seule passation de l'instrument et ne tient pas compte de la stabilité temporelle.³

³ Le Corff Y., Yergeau E., *La méthode de la cohérence interne*, Université de Sherbrooke (Online) : <http://psychometrie.espaceweb.usherbrooke.ca/la-methode-de-la-coherence-interne>

Bien que l'on note une absence de véritable consensus autour de l'interprétation de l'alpha de Cronbach [69], celle qui est souvent retenue est celle de Nunnally [70], [71] ; c'est l'interprétation choisie dans ce travail (voir tableau 3).

Valeur de l'Alpha de Cronbach	Niveau d'accord
< 0,7	Insuffisant
0,70 - 0,80	Satisfaisant
> 0,80	Excellent

Tableau 3 - Interprétation de l'Alpha de Cronbach selon Nunnally

L'**erreur standard de mesure (SEM)** qui permet de quantifier l'écart séparant les mesures répétées : en effet, une erreur de mesure a lieu à chaque fois que l'on utilise des échantillons aléatoires ; la SEM illustre l'importance de cette erreur [64]. Lors d'une évaluation, on retrouvera systématiquement une erreur de mesure, aussi infime soit-elle. Les biais possibles proviennent de trois potentielles sources distinctes : l'évaluateur qui réalise la mesure, le patient évalué et l'outil utilisé [72], [73]. La SEM est utilisée généralement pour calculer le changement minimum détectable, qui est le changement au-delà de l'erreur de mesure, généralement associé à la différence minimale cliniquement importante, qui est le plus petit changement de score que les patients perçoivent comme étant important [64]. Au plus cette SEM sera faible, au plus l'outil sera fiable.

2.2 Méthode de recherche des études

2.2.1 Sources documentaires investiguées

Pour commencer, des recherches rigoureuses ont été effectuées via **PubMed**, **Cochrane**, **DiTA** et **BioMedicalCentral**, où il est possible de trouver des articles traitant de questions diagnostiques comme la nôtre. Cependant, des recherches supplémentaires ont eu lieu sur les bases de données **PEdro** et **ClinicalTrials**, car même si ces bases de données ne présentent pas de protocoles répondant à une question diagnostique en tant que telle, les articles qui y sont présents peuvent tout de même contenir des paragraphes traitant des propriétés clinimétriques de certains outils de mesure. Des recherches complémentaires ont également été menées via **Google Scholar**, mais aussi sur **KineDoc**, cette dernière base de données ayant permis de trouver de la littérature grise que sont les mémoires étudiants, utiles du fait de leurs sources. Certains articles ont également été fournis grâce à des **conseils d'experts**.

2.2.2 Équation de recherche utilisée

N.B. : Ce mémoire, qui étudie principalement la fiabilité des outils de mesure de la proprioception chez le sujet hémiplégique, a été réalisé dans le cadre d'un travail plus vaste en collaboration avec une équipe de Centre de Recherche Interdisciplinaire en Réadaptation (CRIR) de Montréal avec un objectif de publication ; ce travail traite de l'ensemble des paramètres métrologiques des outils d'évaluation de la proprioception chez le sujet hémiplégique. Les équations présentées ci-dessous comprennent donc des paramètres non-abordés dans ce mémoire, comme la validité par exemple. De plus, les tris et extractions ont été effectués en double, par Adrien PENISSON (AP) et Karim JAMAL (KJ) sous le contrôle de Cyril DUCLOS (CD).

L'équation de recherche #5 a été utilisée sur **PubMed** (voir figure 10) moteur de recherche permettant entre autres l'accès aux bases de données MEDLINE et OLDMEDLINE [74], conduisant ainsi à l'apparition de 2325 résultats de recherche :

Search	Actions	Details	Query	Results
#5	...	>	Search: (((((((((proprioception[Text Word]) OR (kinesthe*[Text Word])) OR (movement sensation[Text Word])) OR (somatosensory[Text Word])) OR (proprioception[MeSH Terms]) OR position sense*[Text Word] OR sense of position [Text Word])) OR (sensibility[Text Word])) OR ('sensitivity disorders'[Text Word])) AND (((((((((((((((measure*[Text Word]) OR (assess*[Text Word])) OR (scor*[Text Word])) OR (evaluat*[Text Word])) OR (test*[Text Word])) OR (scale[Text Word])) OR (valid*[Text Word])) OR (reliability[Text Word])) OR (correlation[Text Word])) OR (mdc[Text Word])) OR (psychometric*[Text Word])) OR (clinimetric*[Text Word])) OR (quality[Text Word])) OR (responsiveness[Text Word])) OR (sensibility[Text Word])) OR (feasibility[Text Word])) OR (correlation of data[MeSH Terms])) OR (psychometric*[MeSH Terms])) OR (reproducibility of results[MeSH Terms])))) AND (((((((hemiplegi*[Text Word]) OR (cva[Text Word])) OR (avc[Text Word])) OR (cerebrovascular accident[Text Word])) OR (hemipare*[Text Word])) OR (stroke[Text Word])) OR (stroke[MeSH Terms])) NOT (((((((child[Text Word]) OR (children[Text Word])) OR (childhood[Text Word])) OR (perinatal[Text Word])) OR (review*[Text Word])) OR (metaanalysis[Text Word])) OR (animal*[Text Word]))))	2,325
#4	...	>	Search: (((((((child[Text Word]) OR (children[Text Word])) OR (childhood[Text Word])) OR (perinatal[Text Word])) OR (review*[Text Word])) OR (metaanalysis[Text Word])) OR (animal*[Text Word]))	11,991,997
#3	...	>	Search: (((((((hemiplegi*[Text Word]) OR (cva[Text Word])) OR (avc[Text Word])) OR (cerebrovascular accident[Text Word])) OR (hemipare*[Text Word])) OR (stroke[Text Word])) OR (stroke[MeSH Terms]))	349,394
#2	...	>	Search: (((((((((((((((((((measure*[Text Word]) OR (assess*[Text Word])) OR (scor*[Text Word])) OR (evaluat*[Text Word])) OR (test*[Text Word])) OR (scale[Text Word])) OR (valid*[Text Word])) OR (reliability[Text Word])) OR (correlation[Text Word])) OR (mdc[Text Word])) OR (psychometric*[Text Word])) OR (clinimetric*[Text Word])) OR (quality[Text Word])) OR (responsiveness[Text Word])) OR (sensibility[Text Word])) OR (feasibility[Text Word])) OR (correlation of data[MeSH Terms])) OR (psychometric*[MeSH Terms])) OR (reproducibility of results[MeSH Terms]))	12,114,078
#1	...	>	Search: (((((((((proprioception[Text Word]) OR (kinesthe*[Text Word])) OR (movement sensation[Text Word])) OR (somatosensory[Text Word])) OR (proprioception[MeSH Terms]) OR position sense*[Text Word] OR sense of position [Text Word])) OR (sensibility[Text Word])) OR ('sensitivity disorders'[Text Word]))	83,826

Figure 8 - Recherche PubMed

La base de données **Cochrane** a aussi été investiguée (voir figure 11) avec l'équation suivante sur son manager de recherche avancée, cela a conduit à l'obtention de 389 résultats de recherche :

#1	(Proprioception OR kinesthe* OR Movement sensation OR somatosensory OR position sense OR sense of position OR 'sensitivity disorders');ti,ab,kw	Limits	8137
#2	(measure* OR assess* OR scr* OR evaluat* OR test* OR scale OR Valid* OR reliability OR correlation OR mdc OR psychometric* OR Clinimetric* OR quality OR responsiveness OR sensibility OR feasibility);ti,ab,kw	Limits	1157355
#3	(hemipleg* OR cva OR avc OR cerebrovascular accidnent OR hemipare* OR stroke);ti,ab,kw	Limits	57634
#4	(child* OR perinatal OR review* OR metaanalysis OR animal*);ti,ab,kw	Limits	258651
#5	#1 AND #2 AND #3 NOT #4	Limits	389

Figure 9 - Recherche Cochrane

Également, a été lancée une recherche sur **DiTA**, une base de données produite par PEDro en 2019 [75] et dédiée à la recherche diagnostique. Les mots-clés « STROKE PROPRIO* ASSESS* » ont été utilisés, cependant aucun article pertinent n'a été trouvé, probablement à cause de la récente création du site.

Les recherches sur **BioMedCenter** et **PEDro**, là encore avec les mêmes mots-clés, ont quant à elles conduit à l'enregistrement respectif de 520 et 30 résultats.

Une recherche sur **Google Scholar** a également été effectuée à base des mots-clés précédemment cités, cependant, au vu du grand nombre de résultats retrouvés (33 500), seuls les 30 premiers ont été retenus, en accord avec KJ et CD.

Sur **ClinicalTrials**, 6 résultats sont ressortis de la recherche avancée (voir figure 12).

6 Studies found for: **proprioception assessment | Stroke | Adult, Older Adult | Studies that accept healthy volunteers**

Applied Filters: **Adult (18–64)** **Older Adult (65+)** **Accepts Healthy Volunteers**

Figure 10 - Recherche ClinicalTrials

Tous les synonymes employés ont été trouvés par le biais de connaissances personnelles, des conseils d'experts ainsi que grâce à des recherches effectuées dans le MeSH de l'Inserm⁴.

A cela s'ajoutent 18 études obtenues par **conseils d'experts** et **sources bibliographiques** de littérature.

Au total, ont donc été obtenus aux termes de ces recherches **3318 résultats**.

⁴ Inserm, "Le MeSH bilingue," 2020. [Online]. Available: <http://mesh.inserm.fr/FrenchMesh/>

Enfin, une dernière recherche complémentaire a été effectuée sur le site **Kinedoc** en français (voir figure 13), permettant ainsi d'obtenir 10 résultats de plus, tous des mémoires d'étudiants. Cette littérature grise aura été utile grâce à ses sources.

The screenshot shows a search interface with three rows of search criteria. Each row starts with a dropdown menu labeled 'Titre,sous-titre,résumé' and a search box. The first row has 'avc' in the search box and 'OU accident vasculaire cerebral OU hemiplegi*' in the filter box. The second row has 'evaluation' in the search box and 'OU test OU + -' in the filter box. The third row has 'proprio*' in the search box and 'OU kinesthesi* OU statesthesi*' in the filter box. Below the search boxes is a button labeled 'Recherche simple' with a checkmark and an 'x' icon. At the bottom, a status bar indicates '10 ressource(s) trouvée(s) en 13.418s' and 'Voir la requête effectuée'. A dropdown menu on the right shows 'Tri : pertinence'.

Figure 11 - Recherche Kinedoc

2.2.3 Évaluation de la qualité méthodologique des études sélectionnées

Afin d'évaluer les biais potentiels de chaque étude retenue, il est essentiel d'utiliser des outils évaluant leur qualité méthodologique. Un biais correspond à une erreur systématique non aléatoire et qui va induire une déviation orientée (positive comme négative) de la vérité des résultats et/ou des déductions [76]. L'utilisation d'une échelle adaptée permet ainsi de s'assurer de la qualité méthodologique des articles et donc de la valeur significative de leurs résultats.

L'échelle QAREL de 2012 (voir Annexe II) a été utilisée pour évaluer la qualité méthodologique des études portant sur la fiabilité des méthodes étudiées, puisqu'elle est présentée comme étant l'échelle de référence pour cela [77], [78]. Elle comporte 11 items, dont les réponses peuvent être OUI, NON, INCERTAIN ou NON APPLICABLE.

L'esprit critique de l'auteur de ce travail permettra également de mettre en lumière certains biais potentiels.

2.3 Méthode d'extraction des données et de synthèse des résultats

2.3.1 Méthode de sélection des études

La méthode utilisée a suivi le protocole Cochrane pour la rédaction des revues systématiques [79]. Un premier tri par lecture des titres et résumés a été effectué par deux évaluateurs (AP et KJ) ; en cas de désaccord un troisième évaluateur (CD) intervenait pour trancher. Le deuxième tri par lecture complète des articles s'est effectué de la même façon, par deux évaluateurs et l'intervention d'un troisième en cas de désaccord. Pour les études indisponibles à la lecture complète ou bien dont les données n'étaient pas directement exploitables pour ce travail, un mail était envoyé aux auteurs concernés. Si deux semaines après aucune réponse n'était parvenue, l'article était exclu de la sélection.

2.3.2 Méthode d'extraction des données

Afin d'extraire correctement les données recherchées, il est nécessaire d'effectuer une première lecture complète de chaque article afin de bien en saisir le sens et ne pas passer à côté d'une information essentielle à sa bonne compréhension. Comme recommandé, un tableau à double entrée synthétisant les données importantes pour chacune des différentes études retenues pour ce mémoire a été créé. Les lignes présentent dans ce tableau correspondent aux articles retenus pour l'extraction des données, il y en a une par population étudiée : en effet certains articles utilisent parfois deux populations distinctes dans leurs travaux, les lignes ont donc été dédoublées dans ces cas-là, puisqu'il s'agit de deux populations différentes. Les colonnes quant à elles se portent au nombre de 82.

Les principaux thèmes des colonnes portent sur :

- Les **caractéristiques de la population étudiée**
 - Nombre total et nombre d'hommes et de femmes
 - Âge moyen et écart-type
 - Latéralité des sujets (nombre de gauchers, droits, ambidextres)
 - Latéralité des lésions cérébrales (gauches, droites ou bilatérales), les caractéristiques de l'AVC (ischémique/hémorragique), son antériorité moyenne en mois et son stade moyen (aiguë, subaiguë ou chronique, selon la classification de l'HAS⁵)
- Les **outils évalués** dans l'article ainsi que leurs conditions d'application si nécessaires (yeux ouverts, yeux fermés, etc.)
- Les **parties du corps** sur lesquels ils ont été testés
- Le **nombre de répétitions**, l'**intervalle de temps** entre chacune d'elles et le **nombre d'évaluateurs**
- La **fiabilité**
 - Intra-évaluateur ou inter-évaluateur
 - Avec quel coefficient
 - Résultats selon chaque partie de corps testée
 - Interprétation de ces résultats donnée dans l'article
- La **consistance interne**
 - Résultats
 - Interprétation de ces résultats données dans l'article
- **L'erreur de mesure standard**
 - Résultats

⁵ Haute Autorité de Santé, « Recommandation de bonne pratique » 2012. [Online]. Available: https://www.has-sante.fr/jcms/c_1334330/fr/accident-vasculaire-cerebral-methodes-de-reeducation-de-la-fonction-motrice-chez-l-adulte

- Le **coefficient utilisé**, cette information est donnée plus tôt dans le tableau, mais cette colonne a pour but de simplifier la lecture en cas de recherche d'information rapide lors de l'extraction
- Une colonne « **commentaire** » a également été ajoutée à la toute fin du tableau pour rajouter d'éventuelles informations pertinentes ne pouvant entrer dans les colonnes précédemment expliquées.

Tout comme le tri, l'extraction des données a été effectuée en double, par AP et KJ.

Un dernier point à soulever est la possible présence de spin. Il s'agit d'un embellissement des résultats de la part des auteurs dans la formulation des conclusions [80]. On observe notamment que le spin est généralement plus faible lorsque le score méthodologique est élevé [81]. Afin d'éviter tout biais potentiel lié à des spins dans les articles étudiés, seules les données chiffrées et éventuelles informations complémentaires dans les parties méthode et résultats ont été extraites ; aucune donnée provenant des conclusions des études retenues ne l'a été.

3 – RESULTATS

3.1 Sélection des études

3.1.1 Diagramme de flux

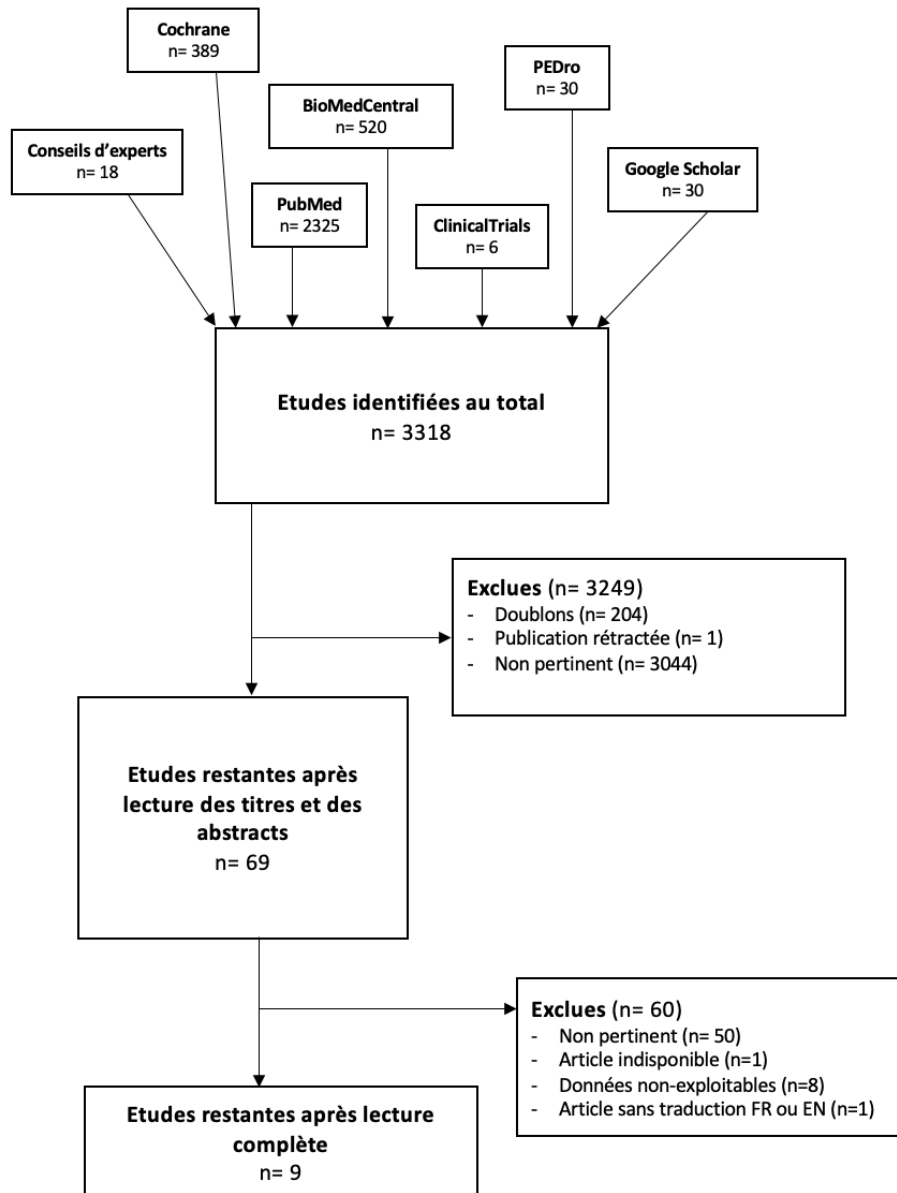


Figure 12 - Diagramme de flux

Après recherche, 3318 articles ont été obtenus : 3300 inclus depuis les différentes bases de données et 18 autres par conseils d'experts et/ou sources bibliographiques.

Après éliminations des doublons (n=204) et des rétractions (n=1), 3249 articles ont été analysés par lecture des titres et résumés : 69 ont été inclus sur la base du respect des critères d'inclusion et d'exclusion. Après lecture complète des articles, 60 ont été exclus, car ils ne traitaient pas du sujet choisi dans ce mémoire (n=50), car ils étaient indisponibles à la lecture

complète (n=1), car les données étaient non-exploitable pour le présent travail (n=8) ou car ils ne disposaient pas de traduction française ou anglaise (n=1).

Finalement, 9 articles ont donc été retenus dans le cadre de ce travail. Cette recherche a suivi le modèle PRISMA [82], [83] et est résumée par un diagramme de flux (voir figure 14).

3.1.2 Études exclues

Le tableau répertoriant les 60 études exclues ainsi que leurs motifs d'exclusion à la suite de la lecture intégrale est joint en *Annexe III*.

3.1.3 Études retenues

9 articles ont été retenus dans cette revue (allant de 1991 à 2020). Il s'agit des suivants :

- **Carey et al. 1996** : Impaired Limb Position Sense After Stroke A Quantitative Test for Clinical Use [84]
- **Gorst et al. 2020** : Assessing lower limb position sense in stroke using the gradient discrimination test (GradDT) and step-height discrimination test (StepDT) : A reliability and validity study [85]
- **Lincoln et al. 1991** : The unreliability of sensory assessments [26]
- **Lincoln et al. 1998** : Reliability and Revision of the Nottingham Sensory Assessment for Stroke Patients [27]
- **Ofek et al. 2019** : Lower extremity position test A new clinical quantitative assessment tool of proprioception post stroke [86]
- **Otaka et al. 2020** : Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke [87]
- **Semrau et al. 2017** : Inter-rater reliability of kinesthetic measurements with the KINARM robotic exoskeleton [88]
- **Sullivan et al. 2011** : Fugl-Meyer assessment of sensorimotor function after stroke : standardized training procedure for clinical practice [89]
- **Villepinte et al. 2018** : Validation of French upper limb Erasmus modified Nottingham Sensory Assessment in stroke [90]

3.2 Extraction des données

Le tableau présenté en *Annexe IV* répertorie l'ensemble des données brutes extraites des 9 études retenues en respectant le modèle évoqué dans la *partie 2.3.2*. Le Tableau 4 ci-dessous présente de façon résumée ces données.

Auteurs et année	Nombre de participants (H/F)	Moyenne d'âge population en années	Moyenne d'antériorité de l'AVC en mois (A/SA/C)	Type d'AVC (H/Isc)	Lésion G/D/BL	Outils étudiés (M/Ins)	Propriétés évaluées	Coefficients utilisés
Carey et al. 1996	35 (28/7)	54	ND	ND	14/21/0	WPST (M)	SEM	NA
Gorst et al. 2020	32 (22/10)	70	22 (C)	7/25	18/14/0	GradDT (Ins) StepDT (Ins)	Fiabilité inter-évaluateur ; Fiabilité intra-évaluateur SEM	ICC
Lincoln et al. 1991	20 (18/2)	ND	< 0,45 (A)	ND/ND	8/11/1	NSA (M)	Fiabilité inter-évaluateur	Kappa
	20 (12/8)		> 1an (C)		4/15/1			
Lincoln et al. 1998	27 (13/14)	ND	> 0,45 (SA)	ND/ND	10/16/1	rNSA (M)	Fiabilité inter-évaluateur	Kappa
Ofek et al. 2019	51 (37/14)	66	52,55 (C)	ND/ND	ND/ND/ND	LEPT (Ins)	Fiabilité intra-évaluateur	ICC
Otaka et al 2020	40 (25/15)	47,6	16,2 (C)	24/16	ND/ND/0	TLT (M)	Fiabilité inter-évaluateur	Kappa
Semrau et al. 2017	15 (12/3)	54,5	2,08 (SA)	3/12	8/5/2	KINARM (R)	Fiabilité inter-évaluateur	ICC
Sullivan et al. 2011	15 (5/10)	62,8	1 (SA)	1/14	9/6/0	FMA	Fiabilité inter-évaluateur ; Fiabilité intra-évaluateur	ICC
Villepinte et al. 2018	50 (35/15)	61,7	3,1 (SA)	13/37	26/24/0	EmNSA	Fiabilité inter-évaluateur ; Fiabilité intra-évaluateur ; Consistance interne	ICC ; Alpha de Cronbach

H Homme - F Femme - A Aigue - SA Subaiguë - C Chronique - He Hémorragique - Isc Ischémique - Ma Manuel - Ins Instrumental - SEM Erreur de mesure standard - NA Non Applicable - ND Non disponible

Tableau 4 - Caractéristiques des études incluses dans la revue

3.3 Analyse des données des études retenues

3.3.1 Qualité méthodologique des articles

Comme expliqué précédemment, l'évaluation des biais potentiels de chaque article a été effectuée à l'aide de l'échelle QAREL (*Annexe II*). Un récapitulatif de cette évaluation, détaillée ci-dessous, est disponible dans le Tableau 5.

L'item 1 pose la question de la représentativité de l'échantillon de population sélectionné par les auteurs ; cette représentativité s'appuie sur un aspect quantitatif (au moins 30 participants nécessaires, voir *Annexe V* pour la justification de cette valeur) et sur un aspect qualitatif (caractéristiques basées sur les données communiquées par l'INSERM⁶ et évoquées en *introduction*). 4 études se sont vues attribuer un NON car elles présentent moins de 30 patients dans leur protocole décrit en méthode [26], [27], [88], [89]. 1 étude s'est vue attribuer un NON car elle présente des patients avec un âge moyen de 25 ans inférieur à l'âge moyen de survenue d'un AVC et qu'elle présente un nombre de patients avec un AVC hémorragique supérieur au nombre de patients présentant un AVC ischémique (sans pour autant que l'AVC hémorragique ou que la précocité de survenue de la pathologie ne soit spécifiquement au cœur de l'étude) [87]. Enfin, 2 études ont été notées avec INCERTAIN car elles ne décrivent pas le type d'AVC dont ont été victimes leurs sujets [84], [86].

L'item 2 pose la question de la représentativité des évaluateurs sélectionnés par les auteurs. 2 études ont été notées avec INCERTAIN : en effet, aucune information quant à la qualification des évaluateurs n'était mentionnée [84], [88].

L'item 3 évalue que les deux assessesurs étaient en aveugles l'un par rapport à l'autre, et concerne uniquement les études traitant de la fiabilité inter-évaluateur, par conséquent 2 études n'étaient pas concernées [84], [86]. 4 études ont été notées avec INCERTAIN, car elles ne donnaient aucune indication à ce sujet dans leur protocole [85], [87]–[89].

L'item 4 évalue si l'assesseur était en aveugle par rapport à son propre résultat précédent et concerne uniquement les études n'employant qu'un seul évaluateur, par conséquent 3 études n'étaient pas concernées [27], [87], [88]. Toutes les autres études ont été notées avec NON, puisque l'évaluation de la proprioception s'effectue par des moyens physiques nécessitant de rencontrer et de dialoguer avec le patient, il est impossible de réaliser cette évaluation en aveugle.

L'item 5 interroge sur le fait que les évaluateurs puissent être en aveugle par rapport au gold standard de la variable étudiée. Du fait du manque de GS évoqué précédemment, cet item était NON APPLICABLE aux études retenues.

Les items 6 et 7 évaluent si les assessesurs étaient en aveugles par rapport à des informations supplémentaires, en lien ou non avec le protocole d'étude. Aucune information à ce sujet n'était disponible dans les articles retenus.

L'item 8 pose la question de la variabilité de l'ordre de passage. Seule 2 étude ont précisé que l'ordre d'évaluation des patients avait varié au cours des différentes sessions

⁶ Inserm, "Accident vasculaire cérébral," 2019. [Online]. Available: <https://www.inserm.fr/information-en-sante/dossiers-information/accident-vasculaire-cerebral-avc>

d'évaluation [27], [84], toutes les autres n'en ont jamais fait mention, sans toutefois préciser que ça n'avait pas été le cas.

L'item 9 demande si l'intervalle de temps entre deux répétitions de mesure est compatible avec la stabilité de la variable mesurée. 3 études se sont vues attribuer un INCERTAIN car l'intervalle évoqué en rapport avec la chronicité de la pathologie était discutable [26], [27], [90]. 1 étude autre étude s'est vu attribuer un INCERTAIN car les auteurs n'ont pas spécifié la taille de cet intervalle [88]. Enfin, 1 étude a obtenu un INCERTAIN car si l'intervalle avait bien été spécifié, ce n'était pas le cas de la chronicité de l'AVC, impossible donc de savoir si cet intervalle était compatible ou non [84].

Les items 10 et 11 portent sur la correcte application/interprétation des outils évalués, ainsi que sur l'utilisation d'une mesure statistique pour juger du niveau d'accord entre les assessesurs. Toutes les études ont répondu positivement à ces critères, grâce aux descriptions d'utilisation des outils présentent dans les articles ainsi qu'à la présence de coefficients calculés.

Auteurs	Année	Outil évalué	01	02	03	04	05	06	07	08	09	10	11	Score total
Carey et al.	1996	WPST	IC	IC	NA	N	NA	IC	IC	O	IC	O	O	3/11
Gorst et al.	2020	GradDT	O	O	IC	N	NA	IC	IC	IC	O	O	O	5/11
Lincoln et al.	1991	NSA	N	O	O	N	NA	IC	IC	IC	IC	O	O	4/11
Lincoln et al.	1998	rNSA	N	O	O	NA	NA	IC	IC	O	IC	O	O	5/11
Ofek et al.	2019	LEPT	IC	O	NA	N	NA	IC	IC	IC	O	O	O	4/11
Otaka et al.	2020	TLT	N	O	IC	NA	NA	IC	IC	IC	O	O	O	4/11
Semrau et al.	2017	KINARM	N	IC	IC	NA	NA	IC	IC	IC	IC	O	O	2/11
Sullivan et al.	2011	FMA	N	O	IC	N	NA	IC	IC	IC	O	O	O	4/11
Villepinte et al.	2018	EmNSA	O	O	O	N	NA	IC	IC	IC	IC	O	O	5/11

O Oui - N Non - IC Incertain - NA Non Applicable

Tableau 5 - Récapitulatif de l'évaluation de la qualité méthodologique des articles par l'échelle QAREL

On observe ainsi que la moyenne obtenue est de 4/11 avec des notes allant de 2/11 à 5/11 ; il s'agit de scores relativement faibles.

3.3.2 Caractéristiques générales de la population

Dans la totalité des 9 études retenues, n=305 sujets atteints d'AVC ont été inclus, parmi lesquels n=207 sont des hommes et n=98 sont des femmes. Parmi eux, n=5 sujets sont gauchers, n=111 sont droitiers, n=0 sont ambidextres (n=189 non déterminés). La moyenne d'âge est de 59,5 ans avec un écart-type moyen de 11,6.

D'après les moyennes obtenues, n=143 sujets évalués sont dans la phase chronique de leur AVC, n=107 dans la phase subaigüe et n=20 dans la phase aigüe (n=35 non déterminés).

La lésion se situe au niveau de l'hémisphère G pour n=97 d'entre eux, D pour n=112 et n=5 ont subi des lésions bilatérales (n=91 non déterminés). Enfin, n=104 sujets ont eu un AVC ischémique et n=48 un AVC hémorragique (n=153 non déterminés).

3.3.3 Outils d'évaluation de la proprioception

Au total, 10 outils de mesure de la proprioception ont été rapportés à travers les 9 études incluses dans cette revue systématique. Parmi ces outils, 5 sont réalisés de façon manuelle et 5 à l'aide d'un instrument (système robotisé, système électronique, tablette avec repères, etc.).

Les outils manuels rapportés sont : la partie proprioception du Nottingham Sensory Assessment (NSA) [26], la partie proprioception de la Revised Nottingham Sensory Assessment (rNSA) [27], le Thumb Localizing Test (TLT) [87], la partie proprioception du Fugl-Meyer Assessment (FMA) [89] et la partie proprioception du Erasmus MC Modifications to the Nottingham Sensory Assessment français (EmNSA) [90].

Les outils instrumentaux rapportés sont quant à eux une version instrumentale du Wrist Position Sense Test (WPST) [84], le GradDT et le StepDT [85], le Lower Extremity Position Test (LEPT) [86] et le KINARM Robot Exoskeleton [88].

La description de tous ces outils est présente en *Annexe VI*.

3.3.4 Articulations évaluées

Au niveau du MS ont été évalués :

Le membre supérieur dans son ensemble, sans distinction d'articulation, par 4 outils [24], [88]–[90] l'épaule, le coude et la main par 2 outils [26], [27] et le poignet par 3 outils [26], [27], [84].

Au niveau du MI, ont été évalués :

Le membre inférieur dans son ensemble pour 2 outils [86], [89], la hanche, le genou et le pied par 2 outils [26], [27] et la cheville par 4 outils [26], [27], [85].

3.3.5 Propriétés métrologiques étudiées

Enfin, parmi les propriétés métrologiques, ont été étudiées :

- La fiabilité inter-évaluateur pour n=8 outils sur 10 [26], [27], [85], [87]–[90]
- La fiabilité intra-évaluateur pour n=6 outils sur 10 [26], [85], [86], [89], [90]
- La consistance interne pour n=1 outil sur 10 [90]
- L'erreur de mesure standard pour n=3 outils sur 10 [84], [85]

3.3.6 Analyse descriptive des données

3.3.6.1 Étude de la fiabilité et de ses composantes (critère de jugement principal)

La **fiabilité inter-évaluateur** a été étudiée pour le GradDT et le StepDT [85], la partie proprioception du NSA [26], la partie proprioception du rNSA [27], le TLT [87], le KINARM [88] et la partie proprioception du FMA [89] et la partie proprioception du EmNSA [90].

Un kappa a été préféré pour la partie proprioception du NSA [26], du rNSA [27] et le TLT [87].

Le NSA a été évalué sur l'épaule, le coude, le poignet, la main, la hanche, le genou, la cheville et le pied de 20 patients. Il a permis d'étudier *l'appréciation du mouvement articulaire*, avec pour résultats : $\kappa = 0,24$ pour l'épaule, $\kappa = 0,19$ pour le coude, $\kappa = 0,35$ pour le poignet, $\kappa = 0,39$ pour la main $\kappa = 0,14$ pour la hanche, $\kappa = 0,14$ pour le genou, $\kappa = 0,28$ pour la cheville et $\kappa = 0,27$ pour le pied. Ces résultats traduisent une mauvaise concordance au niveau de coude, de la hanche et du genou, jusqu'à une faible concordance pour le reste des articulations. Vient ensuite la *sensation de la direction du mouvement* sur ces articulation, obtenant les résultats suivants : $\kappa = 0,16$ pour l'épaule, $\kappa = 0,17$ pour le coude, $\kappa = 0,25$ pour le poignet, $\kappa = 0,26$ pour la main $\kappa = 0,21$ pour la hanche, $\kappa = 0,10$ pour le genou, $\kappa = 0,31$ pour la cheville et $\kappa = 0,31$ pour le pied. Ces valeurs ne correspondent à aucun niveau d'accord pour l'épaule, le coude et le genou, et à un niveau d'accord minimal pour le poignet, la main, la hanche, la cheville et le pied. Le NSA a permis également d'étudier la *sensation de position de l'articulation*, avec $\kappa = 0,06$ pour l'épaule, $\kappa = 0,03$ pour le coude, $\kappa = 0,28$ pour le poignet, $\kappa = 0,30$ pour la main $\kappa = 0,18$ pour la hanche, $\kappa = 0,13$ pour le genou, $\kappa = 0,28$ pour la cheville et $\kappa = 0,28$ pour le pied. Ces valeurs indiquent une fiabilité inter-évaluateur nulle pour l'épaule, le coude, la hanche et le genou et minimale pour le poignet, la main, la cheville et le pied. [26]

Le rNSA a quant à lui été évalué sur les mêmes articulations de 27 patients, cependant les résultats ont été scindés en deux : une évaluation a été réalisée sur les sujets *hémiparétiques gauches* et une autre sur les sujets *hémiparétiques droits*. Pour les hémiparétiques gauches, le kappa obtenu est de $\kappa = 0,51$ pour l'épaule, $\kappa = 0,47$ pour le coude, $\kappa = 0,57$ pour le poignet, $\kappa = 0,32$ pour la main $\kappa = 0,31$ pour la hanche, $\kappa = 0,32$ pour le genou, $\kappa = 0,38$ pour la cheville et $\kappa = 0,46$ pour le pied. Ces valeurs correspondent à un niveau d'accord minimal pour la main, la hanche, le genou et la cheville et à un niveau d'accord faible pour l'épaule, le coude, le poignet et le pied. Enfin, pour les hémiparétiques droits, les résultats sont de $\kappa = 0,47$ pour l'épaule, $\kappa = 0,53$ pour le coude, $\kappa = 0,47$ pour le poignet, $\kappa = 0,32$ pour la main $\kappa = 0,73$ pour la hanche, $\kappa = 0,68$ pour le genou, $\kappa = 0,67$ pour la cheville et $\kappa = 0,62$ pour le pied soit un niveau d'accord minimal au niveau de la main, faible au niveau l'épaule, le coude et le poignet et modéré pour la hanche, le genou, la cheville et le pied. [27]

Il est intéressant de soulever le fait que ces deux études ont été publiées par le même auteur respectivement en 1991 puis en 1998. [26], [27]

Le TLT, évalué sur le MS de 40 patients, a obtenu un kappa de $\kappa = 0,84$ soit un niveau d'accord fort. [87]

C'est l'ICC qui a été choisi par les auteurs pour étudier la fiabilité inter-évaluateur du GradDT [85], du KINARM [88], du FMA [89] et du EmNSA [90].

Le GradDT a été évalué pour la flexion plantaire et la flexion dorsale de cheville de 32 patients. Pour l'exercice proprioceptif de flexion plantaire, l'ICC obtenu est de 0,93 (95%CI = 0,79-0,97) et pour la flexion dorsale 0,92 (95%CI = 0,79-0,97), ce qui indique une excellente fiabilité inter-évaluateur. [85]

Le StepDT évalue le ressenti du patient quant à la hauteur de son pied sur une marche. Ses propriétés ont été évaluées dans les mêmes conditions que celles du GradDT. L'ICC obtenu est ainsi de 0,85 (95%CI = 0,64-0,94), ce qui correspond à une bonne concordance inter-évaluateur. [85]

La partie proprioception du FMA a été évaluée sur le MS et le MI de 15 patients, mais le distinguo n'a pas été effectué dans les résultats. L'ICC ainsi obtenu est de 0,96 avec 95%CI= 0,90-0,99, correspondant ainsi à un excellent niveau de fiabilité inter-évaluateur. [89]

La partie proprioception du NSA a été étudiée sur le MS de 50 patients, et son ICC est de 0,71 avec 95%CI= 0,53-0,83, soit un niveau d'accord moyen. [90]

Enfin, le KINARM a été évalué sur le MS de 15 patients lors de 2 sessions : l'une dans laquelle les patients étaient privés de leur vision et l'autre dans laquelle ils en disposaient. L'ICC correspondant à la session en aveugle est de 0,81 (indiquant un bon niveau d'accord) alors que pour la session avec la vision la valeur obtenue est de 0,95 (soit un excellent niveau d'accord). [88]

L'ensemble de ces résultats est résumé dans le tableau 6.

Auteurs et année	Outil étudié	Coefficient utilisé	Nombre d'évaluateurs	Nombre de sessions (Temps entre chaque session)	Valeur du coefficient 95%CI Interprétation										
					Global	Membre supérieur	Épaule	Coude	Poignet	Main	Membre inférieur	Hanche	Genou	Cheville	Pied
Gorst et al. 2020	GradDT <i>Flexion plantaire</i>	ICC	2	2 (3-7 jours)	ND	ND	ND	ND	ND	ND	ND	ND	ND	0,93* [0,82;0,97] Excellent	ND
	GradDT <i>Flexion dorsale</i>													0,92* [0,79;0,97] Excellent	
	StepDT													0,85* [0,64;0,94] Bon	
Lincoln et al. 1991	NSA <i>Appréciation du mouvement articulaire</i>	Kappa	2	2 (14 jours)	ND	ND	0,24 ND Faible	0,19 ND Mauvais	0,35 ND Faible	0,39 ND Faible	ND	0,14 ND Mauvais	0,14 ND Mauvais	0,28 ND Faible	0,27 ND Faible
	<i>Sensation de direction de mouvement</i>						0,16 ND Mauvais	0,17 ND Mauvais	0,25 ND Faible	0,26 ND Faible		0,21 ND Faible	0,1 ND Mauvais	0,31 ND Faible	0,31 ND Faible
	<i>Sensation de position de l'articulation</i>						0,06 ND Mauvais	0,03 ND Mauvais	0,28 ND Faible	0,3 ND Faible		0,18 ND Mauvais	0,13 ND Mauvais	0,28 ND Faible	0,28 ND Faible
Lincoln et al. 1998	rNSA <i>A gauche</i>	Kappa	2	2 (3-4 jours)	ND	ND	0,51 ND Moyen	0,47 ND Moyen	0,57 ND Moyen	0,32 ND Faible	ND	0,31 ND Faible	0,32 ND Faible	0,38 ND Faible	0,46 ND Moyen

	<i>A droite</i>							0,47 ND Moyen	0,53 ND Moyen	0,47 ND Moyen	0,31 ND Faible		0,73 ND Bon	0,68 ND Bon	0,67 ND Bon	0,62 ND Bon	
Otaka et al. 2020	TLT	Kappa	2	ND	ND	0,84* ND Excellent	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Semrau et al. 2017	KINARM <i>Sans vision</i>	ICC	ND	36 (ND)	ND	0,81* ND Bon	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
	<i>Avec vision</i>					0,95* ND Excellent											
Sullivan et al. 2011	FMA	ICC	2	2 (ND)	0,96 [0,90;0,99] Excellent	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Villepinte et al. 2018	EmNSA	ICC	2	2 (3 jours)	ND	0,71 [0,53;0,83] Moyen	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

ICC Coefficient de Corrélation Intraclasse - ND Non Disponible - * données significatives avec $p < 0,05$ (si pas de * alors non précisé par l'auteur)

Tableau 6 - Résultats des études sur la fiabilité inter-évaluateur

La **fiabilité intra-évaluateur** a été étudiée pour le GradDT et le StepDT [85], la partie proprioception du NSA [26], le LEPT [86], la partie proprioception du FMA [89] et la partie proprioception du EmNSA [90].

Le kappa a été choisi par les auteurs pour évaluer le niveau d'accord autour du NSA, dans les mêmes conditions que précédemment pour la fiabilité inter-évaluateur.

On obtient les résultats suivants pour *l'appréciation du mouvement articulaire* : $\kappa = 0,41$ pour l'épaule, $\kappa = 0,55$ pour le coude, $\kappa = 0,48$ pour le poignet, $\kappa = 0,61$ pour la main $\kappa = 0,57$ pour la hanche, $\kappa = 0,41$ pour le genou, $\kappa = 0,39$ pour la cheville et $\kappa = 0,58$ pour le pied, ce qui correspond à un niveau d'accord faible pour la cheville, moyen pour l'épaule, le coude, le poignet et le reste du MI et une concordance bonne pour la main. Du côté de la *sensation de direction du mouvement* : $\kappa = 0,39$ pour l'épaule, $\kappa = 0,76$ pour le coude, $\kappa = 0,62$ pour le poignet, $\kappa = 0,63$ pour la main $\kappa = 0,71$ pour la hanche, $\kappa = 0,63$ pour le genou, $\kappa = 0,32$ pour la cheville et $\kappa = 0,47$ pour le pied. Ces valeurs correspondent à un niveau d'accord minimal pour l'épaule et la cheville, faible pour le pied et modéré pour le coude, le poignet, la main, le genou et la hanche. Enfin, pour la *sensation de position de l'articulation*, on obtient les chiffres suivants : $\kappa = 0,57$ pour l'épaule, $\kappa = 0,71$ pour le coude, $\kappa = 0,57$ pour le poignet, $\kappa = 0,58$ pour la main $\kappa = 0,38$ pour la hanche, $\kappa = 0,43$ pour le genou, $\kappa = 0,39$ pour la cheville et $\kappa = 0,52$ pour le pied. Ces valeurs correspondent à un niveau d'accord minimal pour la hanche et la cheville, faible pour l'épaule, le poignet, la main, le genou et le pied et modéré pour le coude. [26]

L'ICC a permis de juger de la fiabilité intra-évaluateur du GradDT [85], du LEPT [86], du FMA [89] et du EmNSA [90].

Le GradDT est un outil qui a été employé sur 20 patients post-AVC afin d'évaluer la proprioception de leur cheville dans un premier temps en *flexion plantaire* puis en *flexion dorsale*. Pour la flexion plantaire, l'ICC de la fiabilité intra-évaluateur est de 0,91 avec un 95%CI de 0,82-0,96, signifiant un excellent niveau d'accord. Pour la flexion dorsale, les résultats obtenus sont similaires quoiqu'un peu en deçà, avec un ICC= 0,89 (95%CI de 0,79-0,95) correspondant un bon niveau d'accord. [85]

Le StepDT a été évalué dans les mêmes conditions que le GradDT, obtenant ainsi un ICC=0,95 (95%CI de 0,90-0,97) soit un excellent niveau d'accord. [85]

Le LEPT a été testé sur le MI de 51 patients, avec une déviation de 12 cm (ICC= 0,79) puis de 22 cm (ICC= 0,85), ce qui correspond à un bon niveau de concordance. [86]

La partie proprioception du FMA, testée pour rappel sur le MS et le MI de 15 patients sans distinction dans les résultats, obtient un ICC de 0,95 pour sa fiabilité intra-évaluateur, avec un 95%IC de 0,63-1, soit un excellent niveau d'accord. [89]

La partie proprioception du EmNSA a obtenu un ICC= 0,86 avec un 95%CI de 0,77-0,92, soit un bon niveau d'accord. [90]

L'ensemble de ces résultats est résumé dans le tableau 7.

Auteurs et année	Outil étudié	Coefficient utilisé	Nombre de sessions (Temps entre chaque session)	Valeur du coefficient 95%CI Interprétation												
				Global	Membre supérieur	Épaule	Coude	Poignet	Main	Membre inférieur	Hanche	Genou	Cheville	Pied		
Gorst et al. 2020	GradDT <i>Flexion plantaire</i>	ICC	2 (7-14 jours)	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	0,91* [0,82;0,96] Excellent	ND
	GradDT <i>Flexion dorsale</i>														0,89* [0,79;0,95] Bon	
	StepDT														0,95* [0,90;0,97] Excellent	
Lincoln et al. 1991	NSA <i>Appréciation du mouvement articulaire</i>	Kappa	2 (14 jours)	ND	ND	0,41 ND Moyen	0,55 ND Moyen	0,48 ND Moyen	0,61 ND Bon	ND	0,57 ND Moyen	0,41 ND Moyen	0,39 ND Faible	0,58 ND Moyen		
	<i>Sensation de direction de mouvement</i>					0,39 ND Faible	0,76 ND Bon	0,62 ND Bon	0,63 ND Bon		0,71 ND Bon	0,63 ND Bon	0,32 ND Faible	0,47 ND Moyen		
	<i>Sensation de position de l'articulation</i>					0,57 ND Moyen	0,71 ND Bon	0,57 ND Moyen	0,58 ND Moyen		0,38 ND Faible	0,43 ND Moyen	0,39 ND Faible	0,52 ND Moyen		
Ofek et al. 2019	LEPT <i>Avec 12cm de déviation</i>	ICC	2 (7 jours)	ND	ND	ND	ND	ND	ND	ND	0,79 ND Bon	ND	ND	ND	ND	

	Avec 22cm de déviation										0,85 ND Bon										
Sullivan et al. 2011	FMA	ICC	2 (1 an)	0,95 [0,63;1] Excellent	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	
Villepinte et al. 2018	EmNSA	ICC	2 (7-14 jours)	ND	0,86 [0,77;0,92] Bon	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	

ICC Coefficient de Corrélation Intraclasse - ND Non Disponible - * données significatives avec $p < 0,05$ (si pas de * alors non précisé par l'auteur)

Tableau 7 - Résultats des études sur la fiabilité intra-évaluateur

3.3.6.2 Étude de la consistance interne (critère de jugement secondaire)

La **consistance interne** a été étudiée pour la partie proprioception du EmNSA [90] avec un coefficient de Cronbach.

Le EmNSA obtient un $\alpha = 0,82$ indiquant une consistance interne également satisfaisante.

3.3.6.3 Étude de l'erreur de mesure standard (critère de jugement secondaire)

La **SEM** a été étudiée pour le WPST [84] pour sa fiabilité intra-évaluateur ainsi que le GradDT et le StepDT [85], pour leurs fiabilités inter et intra-évaluateur.

La SEM du WPST a été évaluée à 2 reprises, une première fois entre les sessions d'évaluations 1 et 2 (SEM=2,7°) et entre les sessions d'évaluations 2 et 3 (SEM= 2,9°), soit une moyenne de 2,8°. Les 3 sessions ont été effectuées par le même évaluateur. [84]

Pour la fiabilité inter-évaluateur du GradDT, on retrouve une SEM de 0,45° pour la flexion plantaire et de 0,48° pour la flexion dorsale ; quant à la fiabilité intra-évaluateur, la SEM s'élève à 0,60° pour la flexion plantaire et à 0,63° pour la flexion dorsale. [85]

Enfin, pour la fiabilité inter-évaluateur du StepDT, le résultat obtenu est de 0,30 cm et pour la fiabilité intra évaluateur on obtient une SEM de 0,38cm. [85]

Ces résultats sont résumés dans le tableau 8.

Auteurs et année	Outil étudié	Nombre de sessions	Articulation évaluée	SEM inter-évaluateur	SEM intra-évaluateur
Carey et al. 1996	WPST	3	Poignet	ND	2,8°
Gorst et al. 2020	GradDT Flexion plantaire	2	Cheville	0,45°	0,60°
	GradDT Flexion dorsale			0,48°	0,63°
	StepDT			0,30 cm	0,38 cm

ND Non Disponible

Tableau 8 - Résultats des études sur l'Erreur de Mesure Standard

4 – DISCUSSION

4.1 Analyse et qualité des preuves des principaux résultats

Le principal point négatif mis en lumière par ce travail est le faible nombre d'études se penchant sur la fiabilité des outils d'évaluation de la proprioception chez le patient post-AVC. Cependant, le point positif est que, malgré le faible nombre d'études trouvées, une tendance se dégage : sur les 9 articles, 6 ont moins de 10 ans et 5 ont même moins de 5 ans. Cela pourrait traduire une récente prise de conscience de la part du corps médical et paramédical quant à la pertinence d'utiliser des outils fiables pour effectuer des évaluations complexes. Malheureusement, ces études sont de qualité méthodologique médiocre, comme en témoignent les scores obtenus via l'échelle QAREL, allant de 2 à 5/11 ; de plus, d'autres potentiels biais seront relevés dans cette discussion.

Méthodologiquement, il est nécessaire de proposer d'autres études avec un nombre de patients plus élevé : si nous avons choisi de sélectionner 30 patients comme étant la limite pour que l'échantillon soit représentatif quantitativement parlant comme expliqué dans la *partie 3.3.1*, le nombre théorique de patients à inclure dans une étude sur les AVC est de 207 (*voir Annexe V*), or nous en sommes encore bien loin. Il faut tout de même souligner le fait que ce nombre est très élevé et qu'il est difficile d'inclure autant de patients dans un protocole répondant à nos critères d'inclusion.

Si certaines études distinguent dans leur méthode les patients gauchers ou droitiers, les patients présentant un AVC hémorragique ou ischémique, ou encore la latéralité des lésions, ce n'est jamais le cas dans les résultats. Pourtant, il pourrait être très intéressant d'étudier la fiabilité des outils d'évaluation proprioceptive sur les patients présentant uniquement une hémorragie cérébrale, en opposition aux patients ayant uniquement connu un infarctus cérébral ; ou bien sur les patients droitiers ayant subi des lésions cérébrales dans l'hémisphère gauche en opposition aux droitiers ayant subi des lésions cérébrales dans l'hémisphère droit (et idem pour les patients gauchers) ; ou encore distinguer les résultats selon la phase aiguë, subaiguë ou chronique dans laquelle se situent les patients. Tous ces exemples ne sont que des pistes qu'il semble pertinent d'étudier. Nous pouvons également préciser que la majorité des études ne précisait pas si les patients présentaient ou non des troubles proprioceptifs lors de l'examen, ni si d'autres symptômes pouvaient parasiter leur évaluation (aphasie, etc.), ce qui est regrettable.

4.1.1 La fiabilité inter-évaluateur

D'après les résultats obtenus, le NSA de 1991 et sa version révisée de 1998 ne sont pas des outils fiables. Leur version plus récente, le EmNSA de 2007, présente une fiabilité moyenne, bien que supérieures à ses aînées. Même si les résultats présentés ici ne semblent pas probants, attardons-nous sur certains points. En effet, si d'après ces 3 études la fiabilité inter-évaluateur de la NSA augmente bien au fur et à mesure des modifications qui lui sont apportées, certains points nécessitent tout de même d'être nuancés.

Tout d'abord, l'étude de Lincoln et al de 1991 [26] sur le NSA, qui a obtenu un score de 4/11 sur l'échelle QAREL. Son principal biais est le temps écoulé entre les deux sessions d'évaluation. En effet, il est précisé que les patients ont été évalués pour la première fois durant les deux premières semaines de leur admission en service et que deux autres semaines s'étaient écoulées jusqu'à la deuxième session d'évaluation. A ce stade de l'AVC (phase aiguë), de nombreux progrès ont pu être effectués par les patients durant cet intervalle de temps, d'où peut-être des résultats divergents entre les sujets. En ce qui concerne les évaluateurs et la population de patient, nous pouvons commencer par dire que les évaluateurs présents dans le protocole sont au nombre de 2, il s'agit de deux kinésithérapeutes, l'un étant un kinésithérapeute de service hospitalier et l'autre un kinésithérapeute chercheur. L'échantillon de population est assez faible (20 sujets) et très hétérogène (18 hommes pour 2 femmes), et aucun détail n'est donné quant à la nature de leur AVC (hémorragique ou ischémique) bien que celle-ci tienne un rôle important dans le pronostic de rétablissement des patients. On peut également remarquer que les résultats obtenus au niveau du MS et du MI sont équivalents.

Ensuite, l'étude de Lincoln et al de 1998 [27] sur le rNSA pose question : en effet, 27 patients sont présentés dans cette étude, 10 avec des lésions cérébrales à gauche, 16 des lésions cérébrales à droites et 1 des lésions cérébrales bilatérales. Autrement dit, il y a 10 patients présentant des troubles moteurs et sensitifs à droite et 16 à gauche (donc, pour résumer, 50% de troubles gauches en plus que de troubles droits). Or, les résultats indiquent une bien meilleure concordance à droite qu'à gauche au niveau du membre inférieur. Cela peut sembler logique, car si plus d'individus avec un hémicorps droit sain sont évalués, il sera probablement plus simple pour eux de répondre de façon identique à plusieurs sessions d'un même test, peu importe l'évaluateur. Dire que le rNSA a donc une bonne fiabilité inter-évaluateur au niveau du MI droit est donc peut-être biaisé par le fait qu'il y avait tout simplement plus d'individus aptes à répondre avec précision à l'évaluation de ce membre dans le panel de sujets. Cependant, il est à noter tout de même que si cette différence de résultats gauche/droite est très prononcée au niveau des MI, elle est quasi-inexistante au niveau du membre supérieur, puisque les deux hémicorps obtiennent des résultats extrêmement proches. Outre cet aspect, l'étude a obtenu un score de 5/11 sur l'échelle QAREL. On retrouve un intervalle de temps entre les sessions qui est ici réduit (entre 3 et 4 jours en phase subaiguë). Ici, les évaluateurs présents dans le protocole sont les mêmes que dans l'étude de 1991. L'échantillon de population, bien que plus important (27 sujets) est toujours insuffisant et on retrouve toujours un manque de détails quant à la population et aux caractéristiques des AVC ; cependant ici, le différentiel H/F est faible (13 hommes pour 14 femmes).

Enfin, l'étude de Villepinte [90] est centrée sur le MS et n'aborde pas du tout le niveau d'accord du EmNSA sur le MI. C'est une décision compréhensible que de choisir de se concentrer sur le MS, cependant l'échelle évaluant distinctement les quatre articulations du MS, il aurait probablement été plus pertinent de donner les résultats pour chaque articulation. Dans ce cas précis, il est impossible de savoir à quoi correspond avec exactitude ce résultat, qui traduit une fiabilité inter-évaluateur moyenne, avec un intervalle de confiance assez large.

Les auteurs de l'étude ont été contactés afin de tenter d'apporter des précisions quant à cette problématique, sans succès. L'étude a obtenu un score de 5/11 sur l'échelle QAREL. L'intervalle de temps entre les deux sessions d'évaluation est de 3 jours, ce qui est un délai correct au vu de la phase subaiguë dans laquelle se situent les patients. Le protocole présente les évaluateurs comme étant deux kinésithérapeutes normalement représentatifs de la population de thérapeutes visée. 50 patients sont présentés, ce qui est un échantillon de bonne taille. Même si on pourra reprocher une asymétrie H/F (35 hommes pour 15 femmes), on observe un âge moyen de 61,7 ans, une bonne fréquence H/I des AVC (13 hémorragiques pour 37 ischémiques) ainsi qu'une homogénéité des lésions cérébrales au sein de la population (26 à gauche pour 24 à droite), cette population semble être un échantillon globalement représentatif de la population visée. Lors de l'étude de la fiabilité inter-évaluateur, seuls 46 des 50 patients ont été évalués, sans définir le profil des patients non évalués.

D'après l'étude de Gorst et al. 2020 [85], le GradDT a obtenu d'excellents niveaux d'accord, aussi bien en flexion plantaire qu'en flexion dorsale de cheville avec des intervalles de confiance équivalents et assez réduits. L'étude a été réalisée en 2 sessions, effectuées avec 3-7 jours d'écart sur des patients en phase chronique de leur AVC (22 mois en moyenne). Les évaluateurs employés dans cette étude correspondent à un chercheur (1^{ère} session) et à un kinésithérapeute praticien (2^{ème} session). Il n'est pas précisé si le second évaluateur était en aveugle par rapport aux résultats du premier, cependant on peut noter que c'est le premier évaluateur qui a formé le second à l'utilisation de l'outil GradDT évalué dans ce travail. La population se compose originellement de 32 patients, cependant pour l'étude sur la fiabilité inter-évaluateur, seuls 20 d'entre eux ont été conservés, mais aucune de leurs caractéristiques n'a été communiquée. Il est donc impossible de savoir si les 12 patients écartés étaient plus atteints, s'il s'agissait de patients présentant un AVC ischémique ou hémorragique, ou encore s'il s'agissait des plus âgés. Il aurait sans doute été plus pertinent et transparent de préciser les caractéristiques des 20 patients restant dans l'étude de la fiabilité inter-évaluateur, puisqu'on constate tout de même qu'il y a 37,5% de patients de moins que ceux présentés au départ, ce qui offre une différence significative et peut conduire à des résultats divergents.

Le StepDT quant à lui semble posséder un bon accord inter-évaluateur, associé cependant à un intervalle de confiance large. En parallèle, le protocole étant le même que celui utilisé pour l'étude des propriétés du GradDT, les caractéristiques des évaluateurs et de la population sont les mêmes, ainsi nous faisons face au même problème qu'évoqué précédemment : plus d'un 1/3 des patients annoncés et décrits comme participants à l'étude n'ont pas été évalués dans le cadre de la détermination de la fiabilité inter-évaluateur.

L'évaluation de la fiabilité inter-évaluateur du TLT [87] a conduit à un kappa d'un excellent niveau d'accord entre les évaluateurs. Aucun intervalle de confiance n'est cependant fourni parmi les résultats, mais la valeur-p est inférieure à 0,001, indiquant des résultats normalement significatifs. Le test n'évalue que le MS, et il n'a pas été fait de distinction entre

les résultats obtenus au niveau des différentes articulations. 2 évaluateurs ont été employés, il s'agit de 2 médecins rééducateurs de 6 et 18 ans d'expériences. La population est quant à elle composée de 40 sujets (25 hommes et 15 femmes) d'âge moyen de 54,5 ans, présentant un AVC en phase subaiguë de type ischémique (12) ou hémorragique (3). C'est une population peu représentative du fait de son faible nombre ainsi que son jeune âge.

Le KINARM [88] a été évalué uniquement pour le MS avec et sans la vision. Avec la vision on obtient un bien meilleur résultat en termes de fiabilité inter-évaluateur (ICC= 0,95 soit un excellent niveau d'accord) par rapport à sans vision (ICC= 0,81 soit un bon niveau d'accord). Il faut tout de même rappeler que la plupart des tests de proprioception évoqués évalue la proprioception en privant le patient de sa vision, il est donc raisonnable de penser que c'est dans ces conditions qu'il est pertinent de comparer la fiabilité inter-évaluateur du KINARM avec celle des autres outils. Malheureusement, aucune information sur les évaluateurs n'a été transmise, qu'il s'agisse de leur nombre ou leur qualification. On peut noter comme excellent point le nombre de mesures effectuées qui est conséquent : 36 mesures pour chacun des 15 patients présentés (5 hommes et 10 femmes), dont l'âge moyen est de 54,5 ans ; c'est une population plutôt jeune. L'antériorité moyenne des AVC est de 2 mois, soit une évaluation effectuée en période subaiguë, et il s'agit de 3 AVC hémorragiques pour 12 infarctus cérébraux. Les patients ne présentent pas d'aphasie (critère d'exclusion), et aucune précision n'a été donnée quant à la présence de troubles proprioceptifs : il pourrait être intéressant de comparer les résultats obtenus dans cette étude avec des patients présentant une aphasie afin de voir si cela a un impact dans la bonne utilisation de l'appareil.

L'étude portant sur la partie proprioception du FMA [89] a été effectuée sur les MI et MS de 15 patients. Elle a été menée lors de 2 sessions filmées, permettant d'effacer le biais lié au changement d'état du patient entre les sessions, et a indiqué une excellente concordance. Ce résultat ne fait cependant pas la distinction entre MS et MI, ce qui est regrettable puisque l'échelle fait ne fait pas seulement la distinction entre les deux membres, mais entre les huit articulations, il aurait donc été possible de présenter des résultats amplement plus détaillés. Les évaluateurs se composent d'1 kinésithérapeute dit expert et d'un échantillon de 17 kinésithérapeutes, répartis sur 5 sites cliniques différents aux États-Unis. Les auteurs considèrent que le groupe de 17 kinésithérapeutes forment 1 seule et unique entité, utilisée pour évaluer la fiabilité inter-évaluateur en comparaison avec l'évaluateur expert. Si, en théorie, cela permet une meilleure représentation de la population d'évaluateurs visée, il s'agit d'un choix méthodologique discutable : en effet, bien que formés, ces 17 thérapeutes n'ont peut-être pas les mêmes capacités d'évaluation ou le même ressenti face aux patients. Il aurait probablement été plus pertinent d'effectuer une étude sur la fiabilité inter-évaluateur du FMA entre ces différents kinésithérapeutes, tous formés par la même personne, afin d'évaluer si, à formation égale, le test présentait un bon niveau d'accord. La population se compose quant à elle de 15 patients (5 hommes et 10 femmes), ce qui est un échantillon très faible, avec une moyenne d'âge de 62,8 ans. Il s'agit d'individus

ayant subi un accident vasculaire environ 1 mois avant l'évaluation (phase subaiguë), 14 d'entre eux ayant subi un AVC ischémique et 1 seul un AVC hémorragique.

En bref :

Nous pouvons dire que le NSA semble avoir bien évolué en termes de fiabilité inter-évaluateur au fil de ses modifications. En effet, celle du EmNSA, version la plus récente, est supérieure à celle du rNSA, elle-même supérieure à celle de la version originelle de l'échelle. Cependant, de nombreux facteurs peuvent remettre en question ces résultats, certes logiques, mais qui mériteraient d'être approfondis : population pas toujours représentative, manque de précisions concernant les atteintes vasculaires, manque de précisions dans les résultats ou encore intervalles de temps trop longs entre deux évaluations.

Si le GradDT semble disposer d'un excellent niveau de concordance inter-évaluateur associé à des intervalles de confiances réduits, l'étude qui le présente souffre d'un biais de sélection de population, déjà faible (32 individus), et qui s'est vu amputée de plus d'un tiers sans pour autant préciser le profil des individus restants dans l'étude.

Le StepDT présente de bons résultats mais en deçà du GradDT et avec un intervalle de confiance plus large, associés aux mêmes problèmes que ceux retrouvés précédemment au niveau du biais de sélection.

Le TLT semble prometteur, cependant l'échantillon de population n'est pas assez représentatif et les résultats mériteraient d'être plus détaillés.

Le KINARM présente une excellente fiabilité inter-évaluateur lorsque les patients peuvent s'aider de leur vision pour effectuer le test, cependant elle est moindre lorsqu'ils sont privés de cette dernière. Si l'on compare ce qui est comparable, le KINARM présente donc une fiabilité inter-évaluateur inférieure à certains outils manuels comme le FMA. De plus, bien que le profil des évaluateurs ne soit pas décrit, il est probable que le maniement d'un tel outil robotisé ait nécessité une formation plus ou moins poussée.

Enfin, la partie proprioception du FMA présente une excellente concordance associée à un intervalle de confiance très étroit, cependant la population utilisée dans l'étude est extrêmement réduite et le fait de considérer les 17 évaluateurs non-experts comme une seule entité peut entraîner de nombreux biais.

4.1.2 La fiabilité intra-évaluateur

Pour le NSA, 2 semaines se sont écoulées entre les 2 sessions d'évaluation effectuées, les kappas ainsi obtenus sont globalement meilleurs au niveau du MS que du MI. L'évaluateur présenté est un kinésithérapeute chercheur, et la population, bien que composée de 20 sujets également, est différente de la population utilisée précédemment pour déterminer la fiabilité inter-évaluateur du NSA. Ici, le différentiel H/F est moindre (12 hommes pour 8 femmes) et les sujets sont tous dans une phase chronique de leur AVC car sélectionnés, entre autres, sur le fait que leur accident doit dater d'il y a au moins 1 an. Cependant, une nouvelle fois, aucun détail n'a été donné quant à la nature de l'AVC.

Pour le EmNSA, 3 jours ont séparé les deux évaluations, et les résultats indiquent un bon niveau d'accord au niveau du MS global, sans distinction des articulations et avec un intervalle de confiance correct. L'évaluateur est l'un des deux kinésithérapeutes ayant participé à l'étude de la fiabilité inter-évaluateur. L'échantillon de population étudié est le même que lors de l'évaluation de cette dernière, avec cette fois 48 patients évalués sur les 50 au départ. Au vu des résultats au niveau du MI inférieurs à ceux au niveau du MS dans l'étude d'une version précédente du NSA, il aurait été intéressant d'étudier ici le MI afin de vérifier si la nouvelle version avait pallié ce problème.

Les résultats présentés par l'équipe de chercheurs travaillant sur le GradDT [85] témoignent d'une excellente concordance pour l'évaluation de la proprioception en flexion plantaire de cheville et d'une bonne concordance (dans la tranche haute) pour la flexion dorsale. Les intervalles de confiance là encore équivalents et plutôt réduits, avec une valeur- p inférieure à 0,001, donc des résultats significatifs. Un délai de 7 à 14 jours a été nécessaire entre les deux sessions d'évaluation, ce qui, théoriquement, n'a pas permis un changement d'état significatif des patients, ceux-ci étant dans la phase chronique de leur AVC, qui date d'en moyenne 22 mois. L'évaluateur employé ici est un chercheur, et la population présente les mêmes caractéristiques que présentées en 4.1.1 avec cette fois-ci la totalité des 32 sujets présentés dans le protocole qui a participé à l'étude de la fiabilité intra-évaluateur.

Le StepDT présente une excellente concordance intra-évaluateur associée à un intervalle de confiance très restreint. Les conditions d'évaluation ainsi que les caractéristiques de l'évaluateur et de la population sont les mêmes que citées précédemment.

La fiabilité intra-évaluateur du LEPT a été évaluée avec un intervalle de 7 jours entre les 2 sessions [86]. Que la déviation soit de 12 ou 22cm, cette concordance est interprétée comme étant bonne, quoique légèrement supérieure lors d'une déviation de 22cm. L'ensemble des résultats est donné pour le MI, sans différenciation entre les articulations. Les caractéristiques de l'évaluateur des propriétés métrologiques du LEPT ne sont pas détaillées dans cette étude, et la population comporte 51 sujets (37 hommes pour 14 femmes), âgés en moyenne de 66 ans et avec un AVC en phase chronique (52,55 mois en moyenne). Les caractéristiques des AVC n'ont cependant pas été données.

La partie proprioception du FMA semble présenter une excellente concordance intra-évaluateur, avec cependant un large intervalle de confiance. Si le délai entre les 2 sessions d'évaluation est d'un an, l'état du patient est cependant strictement identique au jour du premier test, et pour cause : la première session a été intégralement filmée, et c'est cet enregistrement qu'a évalué l'assesseur lors de la deuxième session douze mois plus tard, éliminant ainsi le biais que pouvait représenter un changement d'état du patient. L'évaluation a ici été réalisée par un kinésithérapeute expert, avec 30 ans d'expérience. La population présente quant à elle les mêmes caractéristiques que celles évoquées plus haut lors de l'étude de la fiabilité inter-évaluateur.

En bref :

Le NSA et le EmNSA présentent une évolution logique de leur fiabilité intra-évaluateur, à l'image de la fiabilité inter-évaluateur évoquée précédemment. L'EmNSA semble disposer d'une bonne fiabilité intra-évaluateur au niveau du MS, dans la tranche haute de l'interprétation, mais plus de détails concernant les résultats obtenus au niveau de chaque articulation auraient été nécessaires pour effectuer une bonne comparaison.

Le GradDT possède une fiabilité intra-évaluateur très bonne et similaire à la fiabilité inter-évaluateur, cependant l'échantillon de population est plus important et ses caractéristiques plus transparentes.

Le StepDT semble présenter une bien meilleure fiabilité intra-évaluateur qu'inter-évaluateur, avec un biais de sélection qui semble potentiellement moins important.

Le LEPT présente un bon ICC mais la seule étude disponible présente un manque de précision sur les caractéristiques de l'AVC et des résultats.

La partie proprioception du FMA présente un excellent indice d'accord intra-évaluateur, d'autant plus que la deuxième évaluation a été réalisée par le même évaluateur à un an d'intervalle, ce qui a permis à l'assesseur « d'oublier » totalement les notes attribuées aux patients lors de la première session, tout en empêchant tout changement d'état de ces derniers grâce à l'enregistrement vidéo. Il serait cependant intéressant de connaître plus en détails le déroulement de l'évaluation par vidéo, afin de savoir si celle-ci peut être comparable à une évaluation effectuée face au patient en chair et en os.

4.1.3 La consistance interne et l'erreur de mesure standard

Une seule étude s'est penchée sur la consistance interne d'un outil de mesure de la proprioception applicable aux patients post-AVC. La partie proprioception de l'EmNSA (n=4 items) a obtenu une consistance interne satisfaisante [90] ; aucun détail n'a été fourni quant à la qualification ou au nombre d'évaluateurs de cette propriété. On peut cependant noter qu'il s'agit de la consistance interne la plus faible des sous-catégories de la partie somatosensorielle de l'EmNSA, la sensibilité tactile (n=16 items) présentant un $\alpha = 0,95$ et la discrimination de 2 points (n= 2 items) un $\alpha = 0,85$. Il est indiqué qu'en général, au plus le nombre d'items augmente dans une échelle, au plus sa consistance interne diminue⁷ ; nous remarquons que cela ne s'applique pas ici.

La SEM a été étudiée pour le WPST, le GradDT et le StepDT.

Le WPST présente une SEM de fiabilité intra-évaluateur moyenne de 2,8° à savoir 2,7° entre les sessions 1 et 2 et 2,9° entre les sessions 2 et 3. C'est un résultat régulier et qui est bon pour une méthode manuelle évaluée sur 35 patients : en effet, le poignet possède une amplitude de flexion-extension d'environ 180° [91], la SEM calculée ici représente donc une

⁷ J-L. Roulin, "Consistance interne" in *Université Savoie Mont-Blanc*. Updated 2018. [Online]. Available: http://psychometrie.jlroulin.fr/cours/aide_quizz.html?E6334.html

marge d'erreur de 1,5% par rapport à l'amplitude totale physiologique, ce qui semble être tout à fait acceptable et ne pas entraver la bonne prise en charge des patients.

Le GradDT présente une SEM de fiabilité inter-évaluateur de 0,45° pour la flexion plantaire, soit une marge d'erreur de 1,3% par rapport à l'amplitude complète de flexion plantaire, qui s'élève à environ 35° [92]. Pour la flexion dorsale, on retrouve 0,48° soit une marge d'erreur de 1,9% par rapport aux capacités physiologiques moyennes de flexion dorsale, à savoir environ 25° [92]. Ces résultats permettent d'avancer que cette SEM de fiabilité inter-évaluateur est raisonnable, pour les mêmes raisons que citées précédemment pour le WPST. Les résultats de la SEM de fiabilité intra-évaluateur sont de 0,60° pour la flexion plantaire (soit une marge d'erreur de 1,7%) et 0,63 pour la flexion dorsale (soit une marge d'erreur de 2,5%), ce qui, là encore, traduit de bons résultats ; nous pouvons cependant nous interroger sur le fait que la SEM intra-évaluateur soit plus importante que la SEM inter-évaluateur, puisque la logique voudrait que ce soit l'inverse. Pour rappel, la SEM inter-évaluateur a été étudiée auprès de 20 patients, et non les 32 présentés dans le protocole de départ et évalués pour la SEM intra-évaluateur, d'où l'importance de connaître les caractéristiques de ces sujets ; peut-être les 12 sujets écartés présentaient-ils des troubles proprioceptifs plus importants que ceux restants (il ne s'agit que d'un exemple de piste à explorer pour expliquer cette différence entre les résultats inter et intra-évaluateur).

Le StepDT obtient une SEM inter-évaluateur de 0,38cm et une SEM intra-évaluateur de 0,30cm. Pour rappel, comme expliqué en *Annexe VI*, la marche utilisée pour ce test fait 10cm de haut, la SEM représente donc une marge d'erreur de 3,8% pour la fiabilité inter-évaluateur et de 3% pour la fiabilité intra-évaluateur. Si, là encore, ces chiffres sont faibles, et c'est une bonne chose, il faut tout de même souligner qu'ils représentent une marge d'erreur supérieure à ceux précédemment cités. Du reste, la SEM intra-évaluateur est ici inférieure à la SEM inter-évaluateur, ce qui semble logique et remet en question l'hypothèse évoquée ci-dessus concernant le biais de sélection des patients ; cependant, le StepDT n'évaluant pas tout à fait la même composante proprioceptive que le GradDT, cette hypothèse reste toujours valable.

4.2 Niveaux de preuve

Pour évaluer le niveau de preuve de chaque outil, le système GRADE, préconisé dans les protocoles des revues Cochrane⁸, a été utilisé [93], [94]. Il est cependant nécessaire de nuancer son utilisation. En effet, le système GRADE n'est pas une échelle étalonnée, ni le résultat de calculs mathématiques : il s'agit d'une échelle élaborée par de multiples professionnels de la recherche dans un cadre conceptuel et dont la fiabilité est remise en question [95]. Aussi, bien qu'il soit préconisé par les chercheurs des revues Cochrane, il s'agit d'un système relativement utilisateur-dépendant et pas encore tout à fait adapté aux études portant sur la clinimétrie [96] ; certains critères ont donc été jugés de manière subjective, tout

⁸ Cochrane, « *Grade Approach* ». Updated in 2021. [Online]. Available: <https://training.cochrane.org/grade-approach>

en étant contrôlés par des experts et en restant au maximum en accord avec le cadre établi par les auteurs du système GRADE. Par soucis de transparence et pour permettre aux lecteurs de comprendre les niveaux de confiance indiqués pour chaque outil, voici le détail des choix de l'auteur de ce travail, appuyés par la littérature :

- Dans l'obtention du niveau de preuve par le système GRADE, les notes obtenues peuvent être les suivantes [97] :
 - **1/4 très faible niveau de preuve** : l'estimation de l'effet est probablement nettement différente du véritable effet
 - **2/4 faible niveau de preuve** : l'estimation de l'effet peut être nettement différente du véritable effet
 - **3/4 niveau de preuve modéré** : l'estimation de l'effet est probablement proche du véritable effet, mais possiblement nettement différente
 - **4/4 niveau de preuve élevé** : l'estimation de l'effet doit être très proche du véritable effet

- Tous les articles retenus ont suivi un **schéma d'étude transversale**, il est donc admis pour des études clinimétriques que la notation démarre au maximum de points, à savoir 4/4, et se dégrade si nécessaire selon les facteurs influençant la qualité des preuves suivants [64], [96]

- Le **risque de biais** a été jugé subjectivement en rapport avec le score obtenu par l'étude avec l'échelle QAREL :
 - Si l'étude obtenait une note inférieure ou égale à 3, le risque de biais était considéré comme sévère (-2 points)
 - Si l'étude obtenait une note comprise entre 4 et 6 inclus, le risque de biais était considéré comme modéré (-1 point)
 - Si l'étude obtenait une note supérieure ou égale à 7, le risque de biais était considéré comme faible (-0 point)

Une exception peut être relevée pour la fiabilité inter-évaluateur du GradDT et du StepDT : pour cette composante le risque de biais a été considéré comme étant sévère, alors que le score QAREL obtenu par l'étude est de 5/11, car il a été expliqué dans la *partie 4.1.1* que plus d'un tiers des patients présentés dans le protocole n'avaient pas été évalués pour juger de la fiabilité inter-évaluateur de ces 2 outils, ce qui peut induire un biais de sélection potentiellement important selon leurs profils.

- L'**hétérogénéité des résultats** n'a pu être jugée dans ce travail et a été annotée non applicable, car, du fait de la présence d'1 étude pour 1 composante d'un outil, il a été impossible d'effectuer des comparaisons quant à l'hétérogénéité des résultats obtenus. [96]

- Le **caractère direct des données**, qui s'est traduit par l'obtention de la mention Direct (-0 point) pour chacune des composantes évaluées, comme c'est généralement le cas pour les études portant sur le domaine de la rééducation. [64], [96]
- L'**imprécision**, pour laquelle la notation suivante a été subjectivement établie, selon les données présentées dans la *partie 3.3.6* :
 - Sévère (-2 points) : l'étude ne présentait pas d'intervalle de confiance ET pas de valeur-p
 - Modérée (-1 point) : l'étude présentait un intervalle de confiance > 0,2 OU ne présentait pas d'intervalle de confiance mais présentait une valeur-p < 0,05
 - Non détectée (-0 point) : l'étude présentait un intervalle de confiance < 0,2
- Le **biais de publication**, portant sur la suffisante exhaustivité des recherches des auteurs de ce travail sur les composantes métrologiques étudiées de chacun des outils présentés : si une large recherche a été exécutée, celle-ci ne peut cependant pas se prétendre exhaustive au point d'assurer avoir balayé l'ensemble des études pouvant porter sur les dites-composantes des dits-outils ; c'est pour cela que la notation Probable (-1) a été arbitrairement choisie.

Le résumé des niveaux de preuve attribués, allant de très faible à faible, est disponible dans le tableau 9.

Outils étudiés	Propriétés métrologiques étudiées	Nombre de patients (nombre d'études)	Facteurs influençant la qualité des preuves					Niveaux de preuve
			Risque de biais	Hétérogénéité	Caractère direct des données	Imprécision	Biais de publication	
GradDT	Fiabilité inter-évaluateur	20 (1)	Sévère (-2)	Non applicable	Direct	Non détectée	Probable (-1)	1/4 Très faible
StepDT	Fiabilité inter-évaluateur	20 (1)	Sévère (-2)	Non applicable	Direct	Modérée (-1)	Probable (-1)	1/4 Très faible
NSA	Fiabilité inter-évaluateur	20 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
rNSA	Fiabilité inter-évaluateur	27 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
TLT	Fiabilité inter-évaluateur	40 (1)	Modéré (-1)	Non applicable	Direct	Modérée (-1)	Probable (-1)	1/4 Très faible
KINARM	Fiabilité inter-évaluateur	15 (1)	Sévère (-2)	Non applicable	Direct	Modérée (-1)	Probable (-1)	1/4 Très faible
FMA	Fiabilité inter-évaluateur	15 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
EmNSA	Fiabilité inter-évaluateur	50 (1)	Modéré (-1)	Non applicable	Direct	Non détectée	Probable (-1)	2/4 Faible
GradDT	Fiabilité intra-évaluateur	32 (1)	Modéré (-1)	Non applicable	Direct	Non détectée	Probable (-1)	2/4 Faible
StepDT	Fiabilité intra-évaluateur	32 (1)	Modéré (-1)	Non applicable	Direct	Non détectée	Probable (-1)	2/4 Faible

NSA	<i>Fiabilité intra-évaluateur</i>	20 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
LEPT	<i>Fiabilité intra-évaluateur</i>	51 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
FMA	<i>Fiabilité intra-évaluateur</i>	15 (1)	Modéré (-1)	Non applicable	Direct	Modérée (-1)	Probable (-1)	1/4 Très faible
EmNSA	<i>Fiabilité intra-évaluateur</i>	50 (1)	Modéré (-1)	Non applicable	Direct	Modérée (-1)	Probable (-1)	1/4 Très faible
EmNSA	<i>Constance interne</i>	(1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
GradDT	<i>Erreur de mesure standard fiabilité intra-évaluateur</i>	32 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
GradDT	<i>Erreur de mesure standard fiabilité inter-évaluateur</i>	20 (1)	Sévère (-2)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
StepDT	<i>Erreur de mesure standard fiabilité inter-évaluateur</i>	20 (1)	Sévère (-2)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
StepDT	<i>Erreur de mesure standard fiabilité intra-évaluateur</i>	32 (1)	Modéré (-1)	Non applicable	Direct	Sévère (-2)	Probable (-1)	1/4 Très faible
WPST	<i>Erreur de mesure standard fiabilité intra-évaluateur</i>	54 (1)	Sévère (-2)	Non applicable	Direct	Sévère(-2)	Probable (-1)	1/4 Très faible

Tableau 9 - Synthèse des résultats obtenus par la revue et leurs niveaux de preuve analysés par le système GRADE

4.3 Applicabilité des résultats en pratique

Il existe à ce jour peu d'études traitant de la fiabilité des outils de mesure de la proprioception chez les sujets post-AVC alors que de nombreux outils existent et sont parfois utilisés au quotidien par des professionnels afin d'organiser leurs traitements. Malgré la faible qualité des preuves obtenues, essayons de voir quels outils peuvent se démarquer.

Pour la **pratique clinique**, le NSA et le rNSA semblent bien dépassés, et c'est surtout la dernière version en date, le EmNSA, qui peut être intéressante. Malheureusement si celui-ci présente une bonne consistance interne il se trouve que sa fiabilité n'est pas au niveau de celle des autres outils évalués, en particulier en inter-évaluateur. Le TLT quant à lui semble être prometteur, avec une excellente fiabilité inter-évaluateur, mais on pourra reprocher un manque de transparence dans les données obtenues et aucune étude sur la fiabilité intra-évaluateur. L'outil qui se démarque le plus pour des praticiens semble donc être la partie proprioception du FMA : étant donné ses excellents résultats, son vaste champ d'évaluation (la proprioception de chaque articulation est évaluable) et le fait qu'elle puisse être utilisée aussi bien en centre de rééducation qu'en cabinet par divers professionnels.

Le FMA a l'avantage d'être une méthode d'évaluation manuelle de la proprioception et de s'intégrer dans une échelle évaluant d'autres formes de sensibilité ainsi que la motricité, ce qui représente un gain de temps et d'organisation pour les thérapeutes. De plus, c'est un bilan accessible à tous, pouvant être effectué aussi bien en centre qu'en cabinet. C'est un outil bénéficiant d'une excellente fiabilité inter et intra-évaluateur (malgré un intervalle de confiance un peu large concernant cette dernière), cependant, comme pour tous les outils présentés dans ce travail, seul 1 article a étudié sa fiabilité sur la population de patients visée, ce qui est trop peu, d'autant que sa qualité méthodologique ne l'exempte pas de biais potentiels (pour rappel 4/11 sur l'échelle QAREL) et cela se ressent sur le niveau de confiance qui lui est accordé (très faible). Aussi, l'évaluation de sa fiabilité s'est faite avec des patients en phase subaiguë d'AVC, il pourrait donc être pertinent d'effectuer des recherches plus approfondies sur cet outil auprès de patients dans différentes phases de l'atteinte. Il pourra également être pertinent de se pencher sur la validité de cet outil, puisqu'une bonne fiabilité seule ne permet pas d'effectuer un bon bilan.

Le FMA semble donc être l'outil à privilégier compte tenu des résultats à notre disposition, en particulier pour les cliniciens. Si l'échelle présente 5 parties, il est tout à fait possible de réaliser l'évaluation de chacune indépendamment des autres. Cependant, si le reste de l'échelle est fiable et valide, elle présente l'avantage d'évaluer plusieurs aspects de la sensibilité ainsi que de la motricité.

Pour la **recherche**, le GradDT se distingue par ses excellents résultats en termes de fiabilité inter comme intra-évaluateur, avec une SEM très faible. Le StepDT, qui lui est associé, présente également de très bons résultats. Le fait d'utiliser les deux outils simultanément pourrait permettre des recherches approfondies sur le déroulement du pas lors de la marche,

puisque l'un étudie les mouvements de la cheville pendant que l'autre étudie la hauteur du pied. Attention cependant aux conflits d'intérêts possiblement existants, puisque l'équipe ayant conduit l'étude sur la fiabilité de ces outils est la même que celle qui les a conçus. Il pourrait être intéressant de voir publiées d'autres études menées par des équipes indépendantes.

Au niveau du membre supérieur, seul le KINARM est proposé parmi les études présentées ; il s'agit d'un outil extrêmement coûteux, dont l'utilisation nécessite certainement une formation adéquate et avec finalement une fiabilité inter-évaluateur correcte mais pas au niveau du FMA par exemple. Son avantage est probablement sa précision et le fait que tous les résultats puissent être automatiquement informatisés, mais son coût pourra être un frein non négligeable à son utilisation.

4.4 Biais potentiels de la revue

Afin d'évaluer les risques de biais dans ce travail, il est fait référence au score AMSTAR-2, datant de 2017, et qui correspond à une version plus aboutie de la première version du score AMSTAR, en reprenant ainsi 10 des 11 items de cette dernière, pour porter le total à 16 items (voir *Annexe VII*) [98], [99]. Les résultats obtenus sont résumés dans le tableau 10 et détaillés ci-dessous :

- L'item 1 interroge sur le fait que les auteurs aient ou non inclus les critères PICO dans leur question de recherche et dans leurs critères d'inclusion. Cet item est validé par la *partie 2.1* de ce travail.
- L'item 2 interroge sur la réalisation d'un rapport de revue indiquant que la méthode de la revue a été établie avant la rédaction de la dite-revue. Cet item est partiellement validé car un Prospero a été rédigé et déposé sur le site du NIHR⁹, cependant pour ce mémoire nous avons dévié du protocole établi en utilisant notamment d'autres échelles d'évaluation des biais méthodologiques et en ne présentant que les résultats portant sur la fiabilité.
- L'item 3 demande si les auteurs ont expliqué le choix des schémas d'étude à inclure dans la revue, il est validé grâce à la *partie 2.1*.
- L'item 4 interroge sur l'exhaustivité de la recherche documentaire effectuée. Cet item peut être partiellement validé, les recherches ont été conduites dans 8 bases de données différentes notamment à l'aide d'équations appuyées par des experts ; tout est détaillé dans la *partie 2.2*. Cependant, il reste une part d'ombre qui n'a pas pu être exploré, notamment parmi la littérature grise et les protocoles de recherches encore non publiées.

⁹ National Institute for Health Research, Online:
https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=230925

- L'item 5 porte sur la sélection en double des études. Cet item est validé grâce au travail commun de AP et KJ sur la publication prochaine d'un article, dont ce mémoire présente une partie des résultats, comme expliqué dans la *partie 2.2.2*.
- L'item 6 porte l'extraction en double des études. Cet item est validé pour les mêmes raisons que précédemment.
- L'item 7 demande une liste des études exclues ainsi qu'une justification de leur exclusion. Cet item est validé par la *partie 3.1.2*.
- L'item 8 porte sur la description suffisante des études incluses dans la revue. Il est validé puisque tous les détails sont disponibles dans la partie *résultats* ainsi qu'en annexes lorsqu'évoquées.
- L'item 9 interroge sur une suffisante évaluation des risques de biais des études incluses dans la revue. Il peut être validé car les biais méthodologiques ont d'abord été évalués par l'échelle QAREL comme en expliqué dans la *partie 2.2.3*, puis dans la *discussion* d'autres potentiels biais ont été mis en lumière lorsque nécessaire de part un esprit critique.
- L'item 10 demande sur les auteurs ont indiqué les sources de financement des études incluses dans la revue. Cet item n'est pas validé, cette problématique n'a pas été investiguée.
- L'item 13 porte sur la considération des risques de biais lors de l'interprétation des résultats et la discussion. Cet item est validé grâce à l'évaluation du niveau de confiance réalisé avec le système GRADE et présenté dans la *discussion*.
- L'item 14 interroge sur la formulation d'une critique vis-à-vis des résultats présentés dans cette revue. Cet item est validé par la *discussion*.
- L'item 16 porte sur le rapport d'une potentielle source de conflit d'intérêt, financière ou non. Cet item est validé grâce à la *conclusion* de cette revue.
- Les items 11, 12 et 15 portent sur des méta-analyses (analyses quantitatives) ce qui ne correspond pas au travail réalisé ici, qui est une analyse qualitative.

Le score total de cette revue est donc de 10/16 items au total et 10/13 items applicables. Le fait d'avoir pu travailler en collaboration avec une équipe de chercheurs a permis de diminuer considérablement le risque de biais méthodologiques lors de la rédaction de ce mémoire.

Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
Cotation	O	PO	O	PO	O	O	O	O	O	N	NA	NA	O	O	NA	O	11/16

O Oui - PO Partiellement Oui - N Non - NA Non Applicable

Tableau 10 - Évaluation de la qualité méthodologique de la présente revue de littérature

Bien entendu, l'auteur de ce travail ayant lui-même effectué l'évaluation de sa qualité méthodologique, il n'est pas possible de garantir une objectivité totale dans cette évaluation, même si tout a été fait pour s'en rapprocher au maximum et que les résultats ont été validés par un deuxième évaluateur.

5 – CONCLUSION

Il est temps de conclure ce travail. La première chose à retenir est la plus flagrante : malgré une recherche poussée, peu d'articles ont été trouvés et étudiés, ce qui signifie qu'il y a à ce jour peu d'études sur la fiabilité des outils d'évaluation de la proprioception chez les patients post-AVC, alors que l'intérêt d'un bon diagnostic des troubles proprioceptifs a été soulevé à plusieurs reprises dans cette revue. Il est important d'effectuer des bilans de qualité, d'abord pour permettre un meilleur suivi des patients, mais également pour maintenir la crédibilité de la profession et justifier des actes effectués lors de leur prise en charge ; c'est la responsabilité morale, éthique et professionnelle des thérapeutes qui est mise en jeu lors d'un bilan.

Au vu du faible niveau de preuve présenté pour chacun des outils, il est difficile de statuer sur l'outil le plus pertinent à utiliser en termes de fiabilité ; cependant, pour les raisons expliqués précédemment, le Fugl-Meyer Assessment semble être l'outil qui tire son épingle du jeu pour les praticiens.

Le principal point négatif que ce travail a mis en avant est le faible nombre d'études retrouvées traitant du sujet abordé. Malgré cela, comme expliqué précédemment, 6 études sur 9 ont tout de même moins de 10 ans, nous sommes donc en droit d'espérer que la recherche tende à s'orienter de plus en plus sur la fiabilité des outils utilisés lors des bilans. A ce jour, les quelques études retrouvées sont cependant de qualité méthodologique discutable et les résultats ne présentent que de faibles (voire très faibles) niveaux de preuve. Le faible nombre d'études trouvées, en opposition à la multitude d'outils existant, induit également un autre problème : on retrouve schématiquement 1 étude pour 1 échelle, ce qui est insuffisant, surtout au vu des faibles qualités méthodologiques évaluées. Du reste, sur les 6 études les plus récentes, seule la moitié se penche sur la fiabilité d'outils manuels, l'autre moitié étudiant les propriétés de nouveaux outils instrumentaux : les outils manuels sont les plus à même d'être utilisés par des praticiens-cliniciens, et devraient mériter plus d'attention, il ne faut pas avoir peur de remettre en cause les pratiques d'aujourd'hui afin d'améliorer celles de demain. Insistons sur les faibles niveaux de preuve obtenus, avec généralement trop peu de patients présents dans les études ou l'utilisation de données imprécises : il pourrait être très pertinent de distinguer les résultats par articulations, comme certaines l'ont déjà fait, de distinguer également les résultats de patients présentant des AVC hémorragiques de ceux des patients présentant des AVC ischémiques, ou encore d'étudier les troubles proprioceptifs sur les hémipariés du côté dominant comparativement aux hémipariés du côté non-dominant par exemple. Toutes ces pistes ne sont que des exemples et on ne peut qu'espérer que des études plus approfondies et complètes verront le jour dans les prochaines années.

Ceci est valable pour tous les outils évoqués (ou non) dans ce travail : un plus grand nombre d'études, plus approfondies et de meilleure qualité est nécessaire afin d'identifier les

outils les plus adaptés pour les patients. Il ne faut pas perdre de vue que ce sont ces derniers qui représentent la clé de voûte de la pratique médicale et paramédicale.

Pour terminer, il est important de nuancer les résultats obtenus dans cette revue : s'il faut être prudent quant aux excellents résultats obtenus par certains outils, il en va de même pour les résultats médiocres. En effet, rappelons une dernière fois que tous les résultats issus de cette revue présentent de bien bas niveaux de confiance ; il ne faut donc pas mettre de côté certains outils trop tôt.

L'auteur de ce mémoire déclare n'avoir ni conflit, ni lien d'intérêt, financiers ou intellectuels, concernant les données publiées dans ce travail.

BIBLIOGRAPHIE

- [1] Y. Béjot, E. Touzé, A. Jacquin, M. Giroud, and J.-L. Mas, "Épidémiologie Des Accidents Vasculaires Cérébraux," *Médecine/Sciences*, vol. 25, no. 8–9, pp. 727–732, 2009, doi: 10.1051/medsci/2009258-9727.
- [2] L. A. Connell, N. B. Lincoln, and K. A. Radford, "Somatosensory impairment after stroke: Frequency of different deficits and their recovery," *Clin. Rehabil.*, vol. 22, no. 8, pp. 758–767, 2008, doi: 10.1177/0269215508090674.
- [3] H. Lee, H. Kim, M. Ahn, and Y. You, "Effects of proprioception training with exercise imagery on balance ability of stroke patients," *J. Phys. Ther. Sci.*, vol. 27, no. 1, pp. 1–4, 2015, doi: 10.1589/jpts.27.1.
- [4] A.-V. Bruyneel, "Évaluation De La Proprioception Tests De Statesthésie Et Kinesthésie," *EMC - Kinésithérapie - Médecine Phys. - Réadaptation*, vol. 12, no. 16, pp. 1–12, 2016, doi: 10.1016/S1283-0887(16)73532-8.
- [5] A. Prochazka and P. Ellaway, "Sensory systems in the control of movement.," *Compr. Physiol.*, vol. 2, no. 4, pp. 2615–2627, 2012, doi: 10.1002/cphy.c100086.
- [6] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *American Journal of Ophthalmology*, vol. 122, no. 4, pp. 608–609, 1996, doi: 10.1016/s0002-9394(14)72148-8.
- [7] S. M. Highstein, "Anatomy and Physiology of the Central and Peripheral Vestibular System: Overview," in *The Vestibular System*, S. M. Highstein, R. R. Fay, and A. N. Popper, Eds. New York, NY: Springer New York, 2004, pp. 1–10.
- [8] H. J. ten Donkelaar, J. Broman, and P. van Domburg, "The Somatosensory System," in *Clinical Neuroanatomy: Brain Circuitry and Its Disorders*, Cham: Springer International Publishing, 2020, pp. 171–255.
- [9] S. Hendry and S. Hsiao, "The Somatosensory System," *Fundam. Neurosci. Fourth Ed.*, pp. 531–551, 2013, doi: 10.1016/B978-0-12-385870-2.00024-X.
- [10] J. C. Tuthill and E. Azim, "Proprioception," *Curr. Biol.*, vol. 28, no. 5, pp. R194–R203, 2018, doi: 10.1016/j.cub.2018.01.064.
- [11] E. R. Laskowski, K. Newcomer-Aney, and J. Smith, "Proprioception," *Phys. Med. Rehabil. Clin. N. Am.*, 2000.
- [12] M. Hytönen, I. Pyykkö, H. Aalto, and J. Starck, "Postural control and age," *Acta Otolaryngol.*, vol. 113, no. 1–2, pp. 119–122, 1993, doi: 10.3109/00016489309135778.
- [13] J. Roll, "Cognitive sixième sens , ou sens premier ?," *Intellectica*, pp. 49–66, 2003.
- [14] C. Foroglou and M. S. Comparison, "Uhrastructure du fuseau neuro-musculaire chez Fhomme," vol. 37, pp. 19–37, 1973.
- [15] G. Pocock, C. D. Richards, and D. . Richards, "Le système nerveux et les sens : le système somatosensoriel," in *Physiologie Humaine et Physiopathologie : les Fondements de la médecine*, Elsevier M., 2019.
- [16] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neurosciences : à la découverte du cerveau*. 2017.
- [17] E. E. Bunton, W. A. Pitney, T. A. Cappaert, and A. W. Kane, "The role of limb torque, muscle action and proprioception during closed kinetic chain rehabilitation of the lower extremity.," *J. Athl. Train.*, vol. 28, no. 1, pp. 10–20, 1993.
- [18] P. Le Cavorzin, "Neurophysiologie de la fonction proprioceptive et récupération

- postlésionnelle,” *Kinesither Rev*, vol. 12, no. 128–129, pp. 7–14, 2012, doi: 10.1016/j.kine.2012.07.005.
- [19] L. U. Pampa, L. S. Cahill, and L. M. Carey, “Somatosensory assessment and treatment after stroke: An evidence-practice gap,” *Aust. Occup. Ther. J.*, vol. 62, no. 2, pp. 93–104, Apr. 2015, doi: 10.1111/1440-1630.12170.
- [20] C. E. Winward, R. Rehabilitation, P. W. Halligan, and E. Psychology, “Current practice and clinical relevance of somatosensory assessment after stroke,” *Clin. Rehabil.*, vol. 2155, no. 99, pp. 48–55, 1999.
- [21] J. Han, G. Waddington, R. Adams, J. Anson, and Y. Liu, “Assessing proprioception: A critical review of methods,” *J. Sport Heal. Sci.*, vol. 5, no. 1, pp. 80–90, 2016, doi: 10.1016/j.jshs.2014.10.004.
- [22] B. D. Beynnon, S. H. Ryder, L. Konradsen, R. J. Johnson, K. Johnson, and P. A. Renström, “The effect of anterior cruciate ligament trauma and bracing on knee proprioception,” *Am. J. Sports Med.*, vol. 27, no. 2, pp. 150–155, 1999, doi: 10.1177/03635465990270020601.
- [23] L. Olsson, H. Lund, M. Henriksen, H. Rogind, H. Bliddal, and B. Danneskiold-Samsøe, “Test-retest reliability of a knee joint position sense measurement method in sitting and prone position,” *Adv. Physiother.*, vol. 6, no. 1, pp. 37–47, 2004, doi: 10.1080/14038190310009894.
- [24] E. Otaka *et al.*, “Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke,” *PLoS One*, vol. 15, no. 7, p. e0236437, 2020, doi: 10.1371/journal.pone.0236437.
- [25] C. Pilbeam and V. Hood-Moore, “Test–retest reliability of wrist joint position sense in healthy adults in a clinical setting,” *Hand Ther.*, vol. 23, no. 3, pp. 100–109, 2018, doi: 10.1177/1758998318770227.
- [26] N. B. Lincoln, J. L. Crow, J. M. Jackson, G. R. Waters, S. A. Adams, and P. Hodgson, “The unreliability of sensory assessments,” *Clin. Rehabil.*, vol. 5, no. 4, pp. 273–282, 1991, doi: 10.1177/026921559100500403.
- [27] N. B. Lincoln, J. M. Jackson, and S. A. Adams, “Reliability and revision of the Nottingham Sensory Assessment for stroke patients,” *Physiotherapy*, vol. 84, no. 8, pp. 358–365, 1998, doi: 10.1016/S0031-9406(05)61454-X.
- [28] F. Stolk-Hornsveld, J. L. Crow, E. P. Hendriks, R. van der Baan, and B. C. Harmeling-van der Wel, “The Erasmus MC modifications to the (revised) Nottingham Sensory Assessment: A reliable somatosensory assessment measure for patients with intracranial disorders,” *Clin. Rehabil.*, vol. 20, no. 2, pp. 160–172, 2006, doi: 10.1191/0269215506cr932oa.
- [29] A. Fugl-Meyer, L. Jaasko, I. Leyman, S. Olsson, and S. Steglind, “The Post-Stroke Hemiplegic Patient : I. A Method for Evaluation of Physical Performance,” no. max 36, pp. 2–4, 1975.
- [30] Y. W. Hsieh, I. P. Hsueh, Y. T. Chou, C. F. Sheu, C. L. Hsieh, and G. Kwakkel, “Development and validation of a short form of the Fugl-Meyer motor scale in patients with stroke,” *Stroke*, vol. 38, no. 11, pp. 3052–3054, 2007, doi: 10.1161/STROKEAHA.107.490730.
- [31] C. E. Winward, P. W. Halligan, and D. T. Wade, “The Rivermead Assessment of Somatosensory Performance (RASP): standardization and reliability data,” *Clin. Rehabil.*, vol. 16, no. 5, pp. 523–533, Aug. 2002, doi: 10.1191/0269215502cr522oa.
- [32] S. P. Dukelow *et al.*, “Quantitative assessment of limb position sense following

- stroke,” *Neurorehabil. Neural Repair*, vol. 24, no. 2, pp. 178–187, 2010, doi: 10.1177/1545968309345267.
- [33] H. M. Lee, J. J. Liao, C. K. Cheng, C. M. Tan, and J. T. Shih, “Evaluation of shoulder proprioception following muscle fatigue,” *Clin. Biomech.*, vol. 18, no. 9, pp. 843–847, 2003, doi: 10.1016/S0268-0033(03)00151-7.
- [34] N. Leibowitz *et al.*, “Automated measurement of proprioception following stroke,” *Disabil. Rehabil.*, vol. 30, no. 24, pp. 1829–1836, 2008, doi: 10.1080/09638280701640145.
- [35] Collège des Enseignants de Neurologie, “Accidents Vasculaires Cérébraux,” in *Référentiel des Collèges en Neurologie*, Elsevier M., 2016.
- [36] J. D. Easton, “Epidemiology of stroke recurrence,” *Cerebrovasc. Dis.*, vol. 7, pp. 2–4, 1997, doi: 10.1159/000108229.
- [37] J. R. Shiber, E. Fontane, and A. Adewale, “Stroke registry: hemorrhagic vs ischemic strokes,” *Am. J. Emerg. Med.*, vol. 28, no. 3, pp. 331–333, 2010, doi: 10.1016/j.ajem.2008.10.026.
- [38] O. Godefroy *et al.*, “Troubles cognitifs et démences dans les suites d’un accident vasculaire cérébral,” *Cah. l’Annee Gerontol.*, vol. 2, no. 2, pp. 131–135, 2010, doi: 10.1007/s12612-010-0051-7.
- [39] J. Abraham, P. S. S. Rao, S. G. Inbaraj, G. Shetty, and C. J. Jose, “An Epidemiological Study of Hemiplegia due to Stroke in South India,” *Stroke*, vol. 1, no. 6, pp. 477–481, 1970, doi: 10.1161/01.STR.1.6.477.
- [40] R. S. Razafindrasata, N. F. Rasaholiarison, J. Razafimahefa, and A. D. Tehindrazanarivelo, “Profil évolutif des déficits moteurs chez des patients atteints d’accident vasculaire cérébral au service de neurologie du CHU Befelatanana, Antananarivo,” *Med. Sante Trop.*, vol. 27, no. 4, pp. 421–425, 2017, doi: 10.1684/mst.2017.0718.
- [41] A. Ouerdiene, S. Khalbous, H. Essoussi, R. Friaa, M. Kolsi, and M. Haddad, “Le syndrome de Déjerine Roussy ou douleur centrale post-accident vasculaire cérébral : à propos de 27 cas,” *Rev. Neurol. (Paris).*, vol. 176, pp. S68–S69, 2020, doi: 10.1016/j.neurol.2020.01.213.
- [42] K. Daniel, C. D. A. Wolfe, M. A. Busch, and C. Mckevitt, “What are the social consequences of stroke for working-aged adults?: a systematic review,” *Stroke*, vol. 40, no. 6, pp. 431–441, 2009, doi: 10.1161/STROKEAHA.108.534487.
- [43] S. H. Suk *et al.*, “Abdominal obesity and risk of ischemic stroke: The Northern Manhattan Stroke Study,” *Stroke*, vol. 34, no. 7, pp. 1586–1592, 2003, doi: 10.1161/01.STR.0000075294.98582.2F.
- [44] H. J. Chen, C. H. Bai, W. T. Yeh, H. C. Chiu, and W. H. Pan, “Influence of metabolic syndrome and general obesity on the risk of ischemic stroke,” *Stroke*, vol. 37, no. 4, pp. 1060–1064, 2006, doi: 10.1161/01.STR.0000206458.58142.f3.
- [45] D. L. Tirschwell, N. L. Smith, S. R. Heckbert, R. N. Lemaitre, W. T. Longstreth, and B. M. Psaty, “Association of cholesterol with stroke risk varies in stroke subtypes and patient subgroups,” *Neurology*, vol. 63, no. 10, pp. 1868–1875, 2004, doi: 10.1212/01.WNL.0000144282.42222.DA.
- [46] T. Kurth, C. S. Kase, K. Berger, E. S. Schaeffner, J. E. Buring, and J. M. Gaziano, “Smoking and the risk of hemorrhagic stroke in men,” *Stroke*, vol. 34, no. 5, pp. 1151–1155, 2003, doi: 10.1161/01.STR.0000065200.93070.32.
- [47] T. Kurth, C. S. Kase, K. Berger, J. M. Gaziano, N. R. Cook, and J. E. Buring, “Smoking and

- Risk of Hemorrhagic Stroke in Women,” *Stroke*, vol. 34, no. 12, pp. 2792–2795, 2003, doi: 10.1161/01.STR.0000100165.36466.95.
- [48] M. Jimenez *et al.*, “Alcohol consumption and risk of stroke,” *Stroke*, vol. 43, no. 4, pp. 939–945, 2012, doi: 10.1161/STROKEAHA.111.639435.
- [49] D. Woo *et al.*, “Effect of untreated hypertension on hemorrhagic stroke,” *Stroke*, vol. 35, no. 7, pp. 1703–1708, 2004, doi: 10.1161/01.STR.0000130855.70683.c8.
- [50] P. A. Wolf, R. D. Abbott, and W. B. Kannel, “Atrial fibrillation as an independent risk factor for stroke: The framingham study,” *Stroke*, vol. 22, no. 8, pp. 983–988, 1991, doi: 10.1161/01.STR.22.8.983.
- [51] K. Strong, C. Mathers, and R. Bonita, “Preventing stroke: saving lives around the world,” *Lancet Neurol.*, vol. 6, no. 2, pp. 182–187, 2007, doi: 10.1016/S1474-4422(07)70031-5.
- [52] E. D. Vidoni and L. A. Boyd, “Preserved motor learning after stroke is related to the degree of proprioceptive deficit,” *Behav. Brain Funct.*, vol. 5, no. 1, p. 36, 2009, doi: 10.1186/1744-9081-5-36.
- [53] E. Ernst, “A review of stroke rehabilitation and physiotherapy,” *Stroke*, vol. 21, no. 7, pp. 1081–1085, 1990, doi: 10.1161/01.STR.21.7.1081.
- [54] S. Lennon, “Physiotherapy practice in stroke rehabilitation: A survey,” *Disabil. Rehabil.*, vol. 25, no. 9, pp. 455–461, 2003, doi: 10.1080/0963828031000069744.
- [55] Collèges Français des Enseignants universitaires de Médecine Physique et de Réadaptation, “Accidents Vasculaires Cérébraux,” in *Référentiel des Collèges en Médecine Physique et de Réadaptation*, Elsevier M., 2018.
- [56] E. D. Vidoni and L. A. Boyd, “Preserved motor learning after stroke is related to the degree of proprioceptive deficit,” *Behav. Brain Funct.*, vol. 5, p. 36, Aug. 2009, doi: 10.1186/1744-9081-5-36.
- [57] C. Lim, “Multi-sensorimotor training improves proprioception and balance in subacute stroke patients: A randomized controlled pilot trial,” *Front. Neurol.*, vol. 10, no. March, pp. 1–9, 2019, doi: 10.3389/fneur.2019.00157.
- [58] G. Tosi, D. Romano, and A. Maravita, “Mirror Box Training in Hemiplegic Stroke Patients Affects Body Representation,” *Front. Hum. Neurosci.*, vol. 11, p. 617, 2017, doi: 10.3389/fnhum.2017.00617.
- [59] S. I. Kim *et al.*, “Proprioception rehabilitation training system for stroke patients using virtual reality technology,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 4621–4624, 2013, doi: 10.1109/EMBC.2013.6610577.
- [60] D. De Santis *et al.*, “Robot-assisted training of the kinesthetic sense: Enhancing proprioception after stroke,” *Front. Hum. Neurosci.*, vol. 8, no. JAN, pp. 1–12, 2015, doi: 10.3389/fnhum.2014.01037.
- [61] N. Elangovan, I. L. Yeh, J. Holst-Wolf, and J. Konczak, “A robot-assisted sensorimotor training program can improve proprioception and motor function in stroke survivors,” *IEEE Int. Conf. Rehabil. Robot.*, vol. 2019-June, pp. 660–664, 2019, doi: 10.1109/ICORR.2019.8779409.
- [62] M. B. Eriksen and T. F. Frandsen, “The impact of patient, intervention, comparison, outcome (Pico) as a search strategy tool on literature search quality: A systematic review,” *J. Med. Libr. Assoc.*, vol. 106, no. 4, pp. 420–431, 2018, doi: 10.5195/jmla.2018.345.
- [63] C. Drucker-Godard, S. Ehlinger, and C. Grenier, “Validité et fiabilité de la recherche,” in *Méthodes de recherche en management*, 2014, pp. 297–331.

- [64] A. Pallot and et al., *Evidence Based Practice en Rééducation : démarche pour une pratique raisonnée*, Elsevier M. 2019.
- [65] L. B. Mokkink *et al.*, "The COSMIN checklist for evaluating the methodological quality of studies on measurement properties," *Qual. Life Res.*, vol. 63, no. 7, p. 32, 2014.
- [66] C. Elie and I. Colombet, "Méthodes d'estimation de la reproductibilité," *Sang Thromb. Vaiss.*, vol. 23, no. 3, pp. 138–145, 2011, doi: 10.1684/stv.2011.0592.
- [67] T. K. Koo and M. Y. Li, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research," *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155–163, 2016, doi: 10.1016/j.jcm.2016.02.012.
- [68] J. R. Landis and G. G. Koch, "The Measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
- [69] R. A. Peterson, "Une méta-analyse du coefficient alpha de Cronbach," *Rech. Appl. en Mark.*, 2011.
- [70] E. M. Andresen, "Criteria for assessing the tools of disability outcomes research," *Arch. Phys. Med. Rehabil.*, vol. 81, no. 12 SUPPL. 2, pp. 15–20, 2000, doi: 10.1053/apmr.2000.20619.
- [71] J. C. Nunnally, *Psychometric Theory*. New-York: McGraw Hill, 1967.
- [72] P. Piette, "Métrologie appliquée à la kinésithérapie : mesures , tests et bilans , concepts fondamentaux," *EMC - Kinésithérapie - Médecine Phys. - Réadaptation*, vol. 0, no. 16, pp. 1–15, 2016, doi: 10.1016/S1283-0887(16)73672-3.
- [73] J. P. Weir, "Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM," *J. Strength Cond. Res.*, vol. 19, no. 1, pp. 231–240, 2005.
- [74] K. Canese and S. Weis, "PubMed: The bibliographic database," *NCBI Handb.*, no. Md, 2013.
- [75] M. A. Kaizik, M. J. Hancock, and R. D. Herbert, "DiTA: a database of diagnostic test accuracy studies for physiotherapists," *J. Physiother.*, vol. 65, no. 3, pp. 119–120, 2019, doi: 10.1016/j.jphys.2019.05.006.
- [76] M. Porta, *A dictionary of epidemiology*, 6th ed. New-York : Oxford University Press, 2014.
- [77] N. Lucas, N. P. Lucas, P. Macaskill, L. Irwig, and N. Bogduk, "The development of a quality tool for studies of diagnostic reliability (QAREL) The development of a quality appraisal tool for studies of diagnostic reliability (QAREL)," no. March, pp. 1–6, 2016.
- [78] A. Pallot and S. Rostagno, "QAREL : traduction française de l'échelle de qualité méthodologique pour les études diagnostiques de fiabilité," *Kinesithérapie*, vol. 20, no. Xxx, pp. 1–2, 2020, doi: 10.1016/j.kine.2019.12.048.
- [79] J. Higgins *et al.*, "Cochrane Handbook for Systematic Reviews of Intervention," *Cochrane Database of Systematic Reviews*, 2021. [Online]. Available: www.training.cochrane.org/handbook.
- [80] T. A. McGrath, M. D. F. McInnes, N. Van Es, M. M. G. Leeflang, D. A. Korevaar, and P. M. M. Bossuyt, "Overinterpretation of research findings: Evidence of 'spin' in systematic reviews of diagnostic accuracy studies," *Clin. Chem.*, vol. 63, no. 8, pp. 1353–1362, 2017, doi: 10.1373/clinchem.2017.271544.
- [81] D. P. Nascimento, L. O. P. Costa, G. Z. Gonzalez, C. G. Maher, and A. M. Moseley, "Abstracts of Low Back Pain Trials Are Poorly Reported, Contain Spin of Information, and Are Inconsistent With the Full Text: An Overview Study," *Arch. Phys. Med. Rehabil.*, vol. 100, no. 10, pp. 1976–1985.e18, 2019, doi: 10.1016/j.apmr.2019.03.024.
- [82] A. Liberati *et al.*, *The PRISMA statement for reporting systematic reviews and meta-*

analyses of studies that evaluate health care interventions: explanation and elaboration, vol. 62, no. 10. 2009.

- [83] M. Gedda, "Traduction française des lignes directrices STROBE pour l'écriture et la lecture des études observationnelles French translation of the STROBE Reporting Guidelines for writing and reading observational," *Kinésithérapie, la Rev.*, vol. 15, no. 157, pp. 34–38, 2015.
- [84] L. M. Carey, L. E. Oke, and T. A. Matyas, "Impaired limb position sense after stroke: a quantitative test for clinical use.," *Arch. Phys. Med. Rehabil.*, vol. 77, no. 12, pp. 1271–1278, Dec. 1996, doi: 10.1016/s0003-9993(96)90192-6.
- [85] T. Gorst, J. Freeman, K. Yarrow, and J. Marsden, "Assessing lower limb position sense in stroke using the gradient discrimination test (GradDT) and step-height discrimination test (StepDT): a reliability and validity study.," *Disabil. Rehabil.*, vol. 42, no. 15, pp. 2215–2223, Jul. 2020, doi: 10.1080/09638288.2018.1554008.
- [86] H. Ofek, M. Alperin, and Y. Laufer, "Lower Extremity Position Test: A new clinical quantitative assessment tool of proprioception post stroke," *NeuroRehabilitation*, vol. 44, no. 4, pp. 479–484, 2019, doi: 10.3233/NRE-182662.
- [87] E. Otaka *et al.*, "Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke," *PLoS One*, vol. 15, no. 7 July, pp. 1–13, 2020, doi: 10.1371/journal.pone.0236437.
- [88] J. A. Semrau, T. M. Herter, S. H. Scott, and S. P. Dukelow, "Inter-rater reliability of kinesthetic measurements with the KINARM robotic exoskeleton.," *J. Neuroeng. Rehabil.*, vol. 14, no. 1, p. 42, May 2017, doi: 10.1186/s12984-017-0260-z.
- [89] K. J. Sullivan *et al.*, "Fugl-Meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trials.," *Stroke*, vol. 42, no. 2, pp. 427–432, Feb. 2011, doi: 10.1161/STROKEAHA.110.592766.
- [90] C. Villepinte *et al.*, "Validation of French upper limb Erasmus modified Nottingham Sensory Assessment in stroke.," *Ann. Phys. Rehabil. Med.*, vol. 62, no. 1, pp. 35–42, Jan. 2019, doi: 10.1016/j.rehab.2018.03.004.
- [91] A. Kapandji, "Biomécanique du carpe et du poignet," *Ann. Chir. la main*, 1986.
- [92] M. Maestro and B. Ferre, "Anatomie fonctionnelle du pied et de la cheville de l'adulte," *Rev. du Rhum. Monogr.*, vol. 81, no. 2, pp. 61–70, 2014, doi: 10.1016/j.monrhu.2014.03.001.
- [93] G. H. Guyatt, "GRADE: an emerging consensus on rating quality of evidence and strength of recommendations," vol. 336, no. April, pp. 1–3, 2008.
- [94] B. P. Kavanagh, "The GRADE system for rating clinical guidelines," *PLoS Med.*, vol. 6, no. 9, pp. 1–6, 2009, doi: 10.1371/journal.pmed.1000094.
- [95] C. X. Xie and G. C. Machado, "Clinimetrics: Grading of Recommendations, Assessment, Development and Evaluation (GRADE)," *J. Physiother.*, vol. 67, no. 1, p. 66, 2021, doi: 10.1016/j.jphys.2020.07.003.
- [96] S. Singh, S. M. Chang, D. B. Matchar, and E. B. Bass, "Chapter 7: Grading a body of evidence on diagnostic tests," *J. Gen. Intern. Med.*, vol. 27, no. SUPPL.1, pp. 47–55, 2012, doi: 10.1007/s11606-012-2021-9.
- [97] H. Balshem *et al.*, "GRADE guidelines: 3. Rating the quality of evidence," *J. Clin. Epidemiol.*, vol. 64, no. 4, pp. 401–406, 2011, doi: 10.1016/j.jclinepi.2010.07.015.
- [98] B. J. Shea *et al.*, "AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both," *BMJ*, vol. 358, 2017, doi: 10.1136/bmj.j4008.

- [99] A. Pallot and S. Rostagno, “AMSTAR-2 : traduction française de l’échelle de qualité méthodologique pour les revues de littérature systématiques,” *Kinesithérapie*, vol. 20, no. Xxx, pp. 1–2, 2020, doi: 10.1016/j.kine.2019.12.050.

ANNEXES

ANNEXE I – Mail de la présidente de la branche KINARM de BKIN Technology

Bonjour Adrien,

Thank you for your interest in proprioception assessment. You are likely very familiar with the research work by Sean Dukelow at the University of Calgary. He's spent the past decade quantifying sensory deficits especially proprioception and kinesthesia. You can see a short bio on his work here: <https://kinarm.com/clinical-applications/stroke-and-transient-ischemic-attack/> (note many links to his lab). Links to his papers can also be found.

He purchased his Kinarm by a single upfront payment of ~ \$250,000 CAD which is typically what most of our users do. Once in the clinic, that model would likely change to a 3 to 5 year payment plan. (This is important when considering the economics; cost/benefit.) He also licenses our Kinarm Standard Tests, <https://kinarm.com/kinarm-products/kinarm-standard-tests/> which is the "brain" behind the hardware. We sell this product separately because a lot of our customers are basic researchers, neuroscientists, who develop their own behavioural task/protocol with our open development architecture. KST is licensed at \$6,500 CAD/year or \$27,000 for 5 years.

Also I should note that if a subject has grip and arm strength, the End-Point platform is preferred because of its rapid setup. The 'tests' are the same.

I hope this is helpful for your studies. In France there are Kinarms at CNRS Marseille and Paris; there are many in Belgium.

Kind regards,
Anne

Anne Vivian-Scott, MBA PEng

President & CEO

www.kinarm.com

140 Railway St., Kingston, ON Canada K7K 2L9

tf (1) 888-533-4393 X701 | o (1) 613-507-4393 X701 | c (1) 613-329-2804

ANNEXE II – Échelle QAREL

Quality Appraisal of Diagnostic Reliability (QAREL) Checklist

Item	Yes	No	Unclear	N/A
1. Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied? (DEF: 3, 4, 5, 7, 8, 9)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2. Was the test performed by raters who were representative of those to whom the authors intended the results to be applied? (DEF 3, 4, 6, 7, 8, 9)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3. Were raters blinded to the findings of other raters during the study? (DEF 10)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Were raters blinded to their own prior findings of the test under evaluation? (DEF 11)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Were raters blinded to the results of the reference standard for the target disorder (or variable) being evaluated? (DEF 12)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design? (DEF 13)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Were raters blinded to additional cues that were not part of the test? (DEF 14)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8. Was the order of examination varied? (DEF 15, 16)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Was the time interval between repeated measurements compatible with the stability (or theoretical stability) of the variable being measured? (DEF 17)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10. Was the test applied correctly and interpreted appropriately? (DEF 18)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
11. Were appropriate statistical measures of agreement used? (DEF 19, 20, 21)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
TOTAL				

DEF numbers relate to items on the QAREL Data Extraction Form
To access the Data Extraction Form, please go to <http://qarel.org>

ANNEXE III – Tableau d'exclusion après lecture complète des articles

AUTEURS	ANNEE	MOTIF D'EXCLUSION
APHM	2016	Essai clinique en cours
Ballardini et al.	2019	Population étudiée non-AVC
Beaulieu et al.	2020	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Bengtson et al.	2014	Pas d'étude de la fiabilité ou de ses composantes
Block et al.	2019	Pas d'étude de la fiabilité ou de ses composantes
Borstad et al.	2016	Pas d'étude de la fiabilité ou de ses composantes
Braun et al.	2018	Étude centrée sur la fréquence d'utilisation des instruments de mesure
Busse et al.	2008	Pas d'étude de la fiabilité ou de ses composantes
Cappello et al.	2015	Population étudiée non-AVC
Carbonaro et al.	2014	Population étudiée non décrite
Carlsson et al.	2019	Impossible d'extraire les données traitant uniquement de la proprioception
Cherpin et al.	2019	Étude centrée sur la corrélation entre proprioception et motricité
Colombo et al.	2018	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Contu et al.	2017	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Cusmano et al.	2014	Données sur la population cible non exploitables
De Santis et al.	2015	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Deshpande et al.	2010	Population étudiée non-AVC
Dukelow et al.	2010	Données sur la population cible non exploitables
Dukelow et al.	2012	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Duncan et al.	1983	Pas d'extraction possible pour la partie proprioception de l'étude
Giorgiono et al.	2009	Pas d'étude de la proprioception

Herrnstadt et al.	2015	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Hiengkaew et al.	2012	Étude centrée sur l'aspect moteur, sur l'équilibre et la marche
Hsu et al.	2012	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Illman et al.	2020	Population étudiée non-AVC
Ingemanson et al.	2019	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Johansson et al.	2017	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Kattenstroth et al.	2013	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Kenzie et al.	2017	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Kocaman et al.	2020	Essai clinique en cours
Kocaman et al.	2020	Étude centrée sur la marche
Lego et al.	1981	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Leibowitz et al.	2008	Population étudiée non-AVC
Lin	2005	Données sur la proprioception non isolées
Lin et al.	2003	Données sur l'outil cible non-exploitable
Loison et al.	2013	Population étudiée non-AVC
Low Choy et al.	1980	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Meyer et al.	2015	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Meyer et al.	2016	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Miller et al.	2009	Étude centrée sur la sensibilité tactile
Mochizuki et al.	2019	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Mrotek et al.	2017	Pas d'étude de la fiabilité
Olsen et al.	2017	Article non disponible
Rinderknecht et al.	2018	Données non-exploitable
Röijezon et al.	2017	Population étudiée non-AVC
Scalha et al.	2011	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Simmatis et al.	2020	Population étudiée non-AVC

Simo et al.	2011	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Simo et al.	2014	Données sur la population cible non exploitables
Squeri et al.	2011	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Steimann et al.	2012	Article rédigé en Allemand
Stolk-Hornsveld et al.	2004	Données sur la population cible non exploitables
Tognetti et al.	2005	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Tyryshkin et al.	2014	Étude ne traitant pas de la proprioception
Watkins et al.	2020	Population étudiée non-AVC
Williams et al.	2006	Données sur la proprioception non isolées
Winward et al.	2001	Étude de corrélation
Wu et al.	2016	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception
Zbytniewska et al.	2019	Pas d'étude des propriétés métrologiques de l'évaluation de la proprioception

ANNEXE IV – Tableau d'extraction de l'intégralité des données brutes

Titre de l'article	Auteurs/Dates	Population (Nb total)	Population (Nb Hommes)	Population (Nb Femmes)	Population (Âge Moyen)	Population (Âge écart-type)	Date AVC (Moyenne en mois)	Date AVC (Ecart-Type)	Stade AVC (d'après moyenne)	AVC Héorragique	AVC Ischémique	Lésion G	Lésion D	Lésion BL	Population (Gauchers)	Population (Droitières)	Population (Ambidextres)
Impaired Limb Position Sense After Stroke A Quantitative Test for Clinical Use	Carey et al. 1996	35	28	7	54	13	NA	NA	NA	NA	NA	14	21	0	NA	NA	NA
Assessing lower limb position sense in stroke using the gradient discrimination test (GradDT) and step-height discrimination test (StepDT) a reliability and validity study	Gorst et al. 2020	32	22	10	70	9	22	18	C	7	25	18	14	0	NA	NA	NA
The unreliability of sensory assessments INTER-RATER RELIABILITY	Lincoln et al. 1991	20	18	2	NA	NA	0,45	NA	SA	NA	NA	8	11	1	NA	NA	NA
The unreliability of sensory assessments INTRA-RATER RELIABILITY	Lincoln et al. 1991	20	12	8	NA	NA	12	NA	C	NA	NA	4	15	1	NA	NA	NA
Reliability and Revision of the Nottingham Sensory Assessment for Stroke Patients	Lincoln et al. 1998	27	13	14	NA	NA	0,45	NA	SA	NA	NA	10	16	1	NA	NA	NA
Lower extremity position test A new clinical quantitative assessment tool of proprioception post stroke	Ofek et al. 2019	51	37	14	66	10	52,55	62,78	C	NA	NA	NA	NA	NA	2	49	0
Reliability of the thumb localizing test and its validity against quantitative measures with a robotic device in patients with hemiparetic stroke	Otaka et al. 2020	40	25	15	47,6	12,2	16,2	NA	C	24	16	NA	NA	0	NA	NA	NA
Inter-rater reliability of kinesthetic measurements with KINARM robotic exoskeleton	Semrau et al. 2017	15	12	3	54,5	13,6	2,08	2,1	SA	3	12	8	5	2	1	14	0
Fugl-Meyer assessment of sensorimotor function after stroke standardized training procedure for clinical practice	Sullivan et al. 2011	15	5	10	62,8	NA	1	NA	SA	1	14	9	6	0	NA	NA	NA
Validation of French upper limb Erasmus modified Nottingham Sensory Assessment in stroke	Villepinte et al. 2018	50	35	15	61,7	NA	3,1	NA	SA	13	37	26	24	0	2	48	0

Outil et conditions 1	Outil et conditions 2	Outil 1 Manuel (M) ou Instrumental (I)	Outil 2 Manuel (M) ou Instrumental (I)	Partie du corps testée : membre supérieur	PDCT : Epau	PDCT : Coude	PDCT : Poignet	PDCT : Pouce	PDCT : Doigt long	PDCT : Main	PDCT : Membre inférieur	PDCT : Hanche	PDCT : Genou	PDCT : Cheville	PDCT : Hallux	PDCT : Orteil	PDCT : Pied	Nombre de répétitions	Ecart de temps entre chaque répétition	Nombre d'évaluateurs
Wrist Position Test (WPST) Sessions 1-2	Wrist Position Test (WPST) Sessions 2-3	I	I	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3	24 heures	1
GraddT (PF)	GraddT (DF)	I	I	0	0	0	0	0	0	0	0	0	0	1	0	0	0	NA	7-14 jours pour sessions 1-2 et 3-7 jours pour sessions 2-3	2
Nottingham Sensory Assessment (NSA) Direction of movement sense	Nottingham Sensory Assessment (NSA) Joint position sense	M	M	0	1	1	1	0	0	1	0	1	1	1	0	0	1	NA	NA	2
Standardized Manual Assessment of Direction of movement sense	Standardized Manual Assessment of Joint position sense	M	M	0	1	1	1	0	0	1	0	1	1	1	0	0	1	2	2 semaines	1
Revised Nottingham Sensory Assessment (rNSA) kinaesthetic part for left-side	Revised Nottingham Sensory Assessment (rNSA) kinaesthetic part for right-side	M	M	0	1	1	1	0	0	1	0	1	1	1	0	0	1	2	3-4 jours	2
Lower Extremity Position Test (LEPT) with 12cm deviation	Lower Extremity Position Test (LEPT) with 22cm deviation	I	I	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	7 jours	1
Thumb Localizing Test (TLT)	0	M	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	54x2	Quelques minutes/heures	2
KINARM Robotic Exoskeleton without vision	KINARM Robotic Exoskeleton with vision	I	I	1	0	0	0	0	0	0	0	0	0	0	0	0	0	36	NA	NA
Fugl-Meyer Assessment (FMA) Partie proprioception	0	M	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	NA	NA	NA
Erasmus Modified version of the (revised) Nottingham Sensory Assessment (EmNSA) Partie proprioception	0	M	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3-14 jours	3

Fiabilité	Fiabilité (Reliability) inter-rater	Reliability Inter-rater (Kappa)	Reliability Inter-rater (ICC)	Reliability Inter-rater Kappa (Valeur) Générale	Reliability Inter-rater Kappa (valeur) Epaule Outil 1	Reliability Inter-rater Kappa (valeur) Epaule Outil 2	Reliability Inter-rater Kappa (valeur) Coude Outil 1	Reliability Inter-rater Kappa (valeur) Coude Outil 2	Reliability Inter-rater Kappa (valeur) Poignet Outil 1	Reliability Inter-rater Kappa (valeur) Poignet Outil 2	Reliability Inter-rater Kappa (valeur) Main Outil 1	Reliability Inter-rater Kappa (valeur) Main Outil 2
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0,16	0,06	0,17	0,03	0,25	0,28	0,26	0,3
1	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0,51	0,47	0,47	0,53	0,57	0,47	0,32	0,32
1	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0,84	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0

Reliability Inter-rater Kappa (valeur) Hanche Outil 1	Reliability Inter-rater Kappa (valeur) Hanche Outil 2	Reliability Inter-rater Kappa (valeur) Genou Outil 1	Reliability Inter-rater Kappa (valeur) Genou Outil 2	Reliability Inter-rater Kappa (valeur) Cheville Outil 1	Reliability Inter-rater Kappa (valeur) Cheville Outil 2	Reliability Inter-rater Kappa (valeur) Pied Outil 1	Reliability Inter-rater Kappa (valeur) Pied Outil 2	Reliability Inter Rater ICC Outil 1 (Valeur)	Reliability Inter Rater ICC Outil 1 (95% CI)	Reliability Inter Rater ICC Outil 2 (Valeur)	Reliability Inter Rater ICC Outil 2 (95% CI)	Reliability inter-rater (Appréciation) Outil 1	Reliability inter-rater (Appréciation) Outil 2
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0,93	0,82-0,97	0,92	0,79-0,97	Excellent	Excellent
0,21	0,18	0,1	0,13	0,31	0,28	0,31	0,28	0	0	0	0	Pauvre	Pauvre
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,31	0,73	0,32	0,68	0,38	0,67	0,46	0,62	0	0	0	0	NA	NA
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	High	0
0	0	0	0	0	0	0	0	0,81	0	0,95	0	High	High
0	0	0	0	0	0	0	0	0,96	0,90-0,99	0	0	High	0
0	0	0	0	0	0	0	0	0,71	0,53-0,83	0	0	Moderate	0

Reliability Intra-rater	Reliability Intra-rater (Kappa)	Reliability Intra-rater Kappa (valeur) Epaule Outil 1	Reliability Intra-rater Kappa (valeur) Epaule Outil 2	Reliability Intra-rater Kappa (valeur) Coude Outil 1	Reliability Intra-rater Kappa (valeur) Coude Outil 2	Reliability Intra-rater Kappa (valeur) Poignet Outil 1	Reliability Intra-rater Kappa (valeur) Poignet Outil 2	Reliability Intra-rater Kappa (valeur) Main Outil 1	Reliability Intra-rater Kappa (valeur) Main Outil 2
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	1	0,39	0,57	0,76	0,71	0,62	0,57	0,63	0,58
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0

Reliability Intra-rater Kappa (valeur) Hanche Outil 1	Reliability Intra-rater Kappa (valeur) Hanche Outil 2	Reliability Intra-rater Kappa (valeur) Genou Outil 1	Reliability Intra-rater Kappa (valeur) Genou Outil 2	Reliability Intra-rater Kappa (valeur) Cheville Outil 1	Reliability Intra-rater Kappa (valeur) Cheville Outil 2	Reliability Intra-rater Kappa (valeur) Pied Outil 1	Reliability Intra-rater Kappa (valeur) Pied Outil 2	Reliability Intra-rater (ICC)	Reliability Intra-rater ICC (Valeur) Outil 1.1	Reliability Intra-rater ICC (Valeur) Outil 1.2	Reliability Intra-rater ICC (95%IC) Outil 1	Reliability Intra-rater ICC (Valeur) Outil 2	Reliability Intra-rater ICC (95% CI) Outil 2	Reliability Intra-rater Appréciation Outil 1	Reliability Intra-rater Appréciation Outil 2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0,91	0	0,82-0,96	0,89	0,79-0,95	Excellent	Excellent
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0,71	0,38	0,63	0,43	0,32	0,39	0,47	0,52	0	0	0	0	0	0	Poor-Moderate	Poor-Moderate
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0,79	0	0	0,85	0	Good	Good
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0,95	0	0,63-1	0	0	High	0
0	0	0	0	0	0	0	0	1	0,86	0	0,77-0,92	0	0	Excellent	0

Consistence Interne Outil 1	Consistence Interne Outil 2	Consistence Interne Cronbach's alpha (Valeur) Outil 1	Consistence Interne Cronbach's alpha (Valeur) Outil 2	Consistence Interne Outil 1 (Appréciation)	Consistence Interne Outil 2 (Appréciation)	SEM Intra-rater reliability Outil 1	SEM Intra-rater reliability Outil 2	SEM Intra-rater Outil 1 95%CI	SEM Intra-rater Outil 2 95%CI	SEM Inter-rater reliability Outil 1	SEM Inter-rater reliability Outil 2	SEM Inter-rater Outil 1 95%CI	SEM Inter-rater Outil 2 95%CI	Coefficients
0	0	0	0	0	0	2,7	2,9	±5,5	±6	0	0	0	0	NA
0	0	0	0	0	0	0,6	0,63	0	0	0,45	0,48	0	0	ICC (reliability)
0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kappa
0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kappa
0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kappa
0	0	0	0	0	0	0	0	0	0	0	0	0	0	ICC
0	0	0	0	0	0	0	0	0	0	0	0	0	0	Kappa
0	0	0	0	0	0	0	0	0	0	0	0	0	0	ICC
0	0	0	0	0	0	0	0	0	0	0	0	0	0	ICC
1	0	0,82	0	Excellent	0	0	0	0	0	0	0	0	0	ICC (reliability), Cronbach's alpha (consistence interne)

ANNEXE V – Détermination de la taille de l'échantillon nécessaire

La taille d'un échantillon représentatif d'une population donnée doit être déterminée par le biais de ce calcul :

$$n = t^2 * p * (1-p) / m^2$$

n : taille de l'échantillon minimale requis

t : niveau de confiance (pour 95% la valeur type est de 1,96)

p : proportion estimée de la population qui présente la caractéristique

m : marge d'erreur (5% étant la valeur admise, soit 0,05)

D'après l'étude de Béjot et al. 2009 portant sur l'épidémiologie des AVC à travers 23 pays, l'incidence des AVC au niveau mondial est de 113 à 410 pour 100 000 habitants par an. Nous prenons la valeur médiane de ces résultats, soit 160 cas pour 100 000 habitants, soit une incidence de 0,16%. Afin de simplifier notre calcul nous considérons que l'incidence correspond à la valeur-p.

On obtient donc :

$$\begin{aligned} &= 1,96^2 * 0,16 * (1 - 0,16) / 0,05^2 \\ &= 3,8416 * 0,16 * (0,84/0,0025) \\ &= 3,8416 * 0,16 * 336 \\ &= \mathbf{206,52} \end{aligned}$$

Il serait donc nécessaire, pour chaque étude, de présenter plus de 200 patients post-AVC pour qu'elle soit quantitativement représentative. Bien sûr, ce nombre de participants est rarement atteignable dans une étude clinimétrique sur des sujets présentant une telle pathologie.

Or, il s'avère qu'une convention existait dans le monde de la recherche. En effet, on considère généralement qu'un minimum de 30 sujets est un échantillon représentatif lors d'une étude, car cela permet le passage d'une analyse par la loi normale à une analyse paramétrique, plus significative.

C'est ainsi que le nombre 30 a été retenu dans ce travail pour juger de la limite de la représentativité quantitative d'un échantillon de population.

ANNEXE VI – Description des 10 outils de mesure présentés dans les études retenues

Les méthodes manuelles :

1- La partie proprioception du **Nottingham Sensory Assessment** :

Sensory assessment form

Name _____ Key: 2 = Normal Examiner _____
 Date of stroke _____ 1 = Impaired Date _____
 Side of body affected: RIGHT / LEFT / NEITHER 0 = Absent
 9 = Unable to test
 Pre-existing disease: YES / NO If YES, what? _____ Oedema present: YES / NO If YES, where? _____

Regions of body	TACTILE SENSATION												KINAESTHETIC SENSATION							
	Temperature		Light touch		Pressure		Pin prick		Tactile localization		Bilateral simultaneous touch		2-point discrimination		Appreciate joint movement		Direction movement sense		Joint position sense	
	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R
Face																				
Trunk																				
Shoulder																				
Elbow																				
Wrist																				
Hand: palm																				
Hand: finger tip																				
Hip																				
Knee																				
Ankle																				
Foot																				

Stereognosis: _____

L	R

Comments: _____

10p COIN BIRO COMB SPONGE CUP
 2p COIN PENCIL SCISSORS FLANNEL GLASS

Kinaesthetic sensations

All three aspects are tested simultaneously: appreciation of movement, its direction and accurate joint position sense. The limb on the affected side of the body is supported and moved by the examiner in various directions but movement is only at one joint at a time. The patient is asked to mirror the change of movement with the other limb. If they cannot do this they are asked to indicate whether a movement has taken place. Three practice movements are allowed prior to the blindfolding. The reverse procedure supporting and moving the unaffected arm is attempted if there is a good recovery of movement in the affected limb.

Appreciation of movement taking place – Patient indicates on each occasion that a movement takes place but the direction is incorrect.

Direction of movement sense – Patient is able to appreciate and mirror the direction of the test movement taking place each time, but is

Joint position sense – Accurately mirrors the test movement to within 10° of the new test position.

Scoring for each modality

- 2 Normal – Correct on all three occasions
- 1 Impaired – Correct on some of the three occasions.
- 0 Absent – Incorrect on all three occasions.

Voir également description *partie 1.1.4.2.*

2- La partie proprioception du **Revised Nottingham Sensory Assessment** :

NOTTINGHAM SENSORY ASSESSMENT										
Name					Examiner					
Date of stroke					Date of assessment					
Side of body affected: RIGHT/LEFT/NEITHER/BOTH										
If BOTH, side tested										
Oedema present: YES/NO					If YES, where					
TACTILE SENSATION							KINAESTHETIC SENSATION			
Regions of the body	Light touch L R		Pressure	Pinprick	Temperature L R		Tactile localisation	Bilateral simultaneous touch	Left	Right
Face										
Trunk										
Shoulder										
Elbow										
Wrist										
Hand										
Hip										
Knee										
Ankle										
Foot										

Kinaesthetic Sensations

All three aspects of movement are tested: appreciation of movement, its direction and accurate joint position sense are assessed simultaneously. The limb on the affected side of the body is supported and moved by the examiner in various directions but movement is only at one joint at a time. The patient is asked to mirror the change of movement with the other limb. Three practice movements are allowed before the blindfolding. The reverse procedure, supporting and moving the unaffected arm, is attempted if there is a good recovery of movement in the affected limb.

The upper limb is tested in sitting and the lower, lying supine.

If hand and wrist score 3 give 3s for elbow, shoulder and trunk.

If ankle and foot score 3 give 3s for knee and hip.

Scoring

- 0 Absent* No appreciation of movement taking place.
- 1 Appreciation of movement taking place* Patient indicates on each occasion that a movement takes place but the direction is incorrect.
- 2 Direction of movement sense* Patient is able to appreciate and mirror the direction of the test movement taking place each time, but is inaccurate in its new position.
- 3 Joint position sense* Accurately mirrors the test movement to within 10° of the new test position.
- 9 Unable to test*

Equipment required

Blindfold.

Voir également description *partie 1.1.4.2.*

3- La partie proprioception de l'Erasmus MC Modifications to the Nottingham Sensory Assessment :

**ERASMUS MC MODIFICATIONS TO THE (REVISED)
NOTTINGHAM SENSORY ASSESSMENT (EmNSA)**

PATIENT NAME :
DATE OF BIRTH :

DATE:		RIGHT				LEFT			
		LT	PR	PP	SB	LT	PR	PP	SB
TACTILE SENSATION & SHARP-BLUNT DISCRIMINATION	UPPER ARM								
	FOREARM								
	HAND								
	FINGERS								
	TOTAL SCORE UE								
	THIGH								
	LEG								
	FOOT								
PROPRIOCEPTION	TOES								
	TOTAL SCORE LE								
	SHOULDER								
	ELBOW								
	WRIST								
	FINGERS								
	TOTAL SCORE UE								
	HIP								
KNEE									
ANKLE									
TOES									
TOTAL SCORE LE									

LT = light touch PR = Pressure PP = pinprick SB = sharp-blunt discrimination
UE = upper extremity LE = lower extremity
0 = absent 1 = impaired 2 = normal

Proprioception

Specified passive movements are tested in only one joint at a time. The starting positions, specific hand grips for the physiotherapist to use, along with the directions of the movement to be tested are described in Appendix 3. The large joints (hip and knee, shoulder and elbow) are moved through approximately a quarter of their total range of motion (ROM). The other joints (wrist and fingers, ankle and toes) are moved throughout the full available range of movement. To demonstrate the procedure, three practice movements are allowed (with the patient's eyes open.) Each joint is then moved three times. The patient is asked, using specific questions, to indicate verbally or non-verbally the direction of the movement taking place. If the patient is incapable of doing this, he is then asked to identify (verbally or non-verbally) when movement is taking place.

- 0 Absent Patient does not detect the movement taking place.
- 1 Impaired Patient detects the movement taking place but the direction is not correct on all three occasions.
- 2 Normal Patient correctly detects the direction of the movement taking place on all three occasions.

Voir également description *partie 1.1.4.2.*

4- La partie proprioception du **Fugl-Meyer Assessment** :

Voir description complète *partie 1.1.4.2.*

5- Le **Thumb Localizing Test** :

Les yeux du patient étant fermés, l'examineur positionne l'un des membres supérieurs du patient et lui demande de venir pincer le pouce de ce membre avec le pouce et l'index opposés.

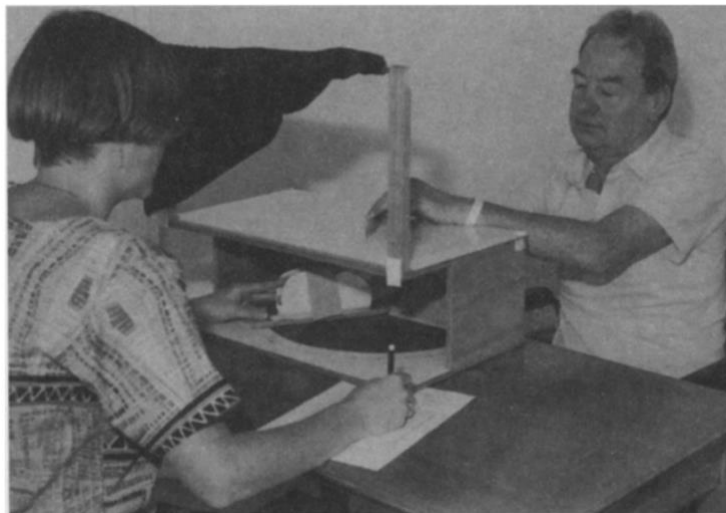
Les méthodes instrumentales :

6- Le **KINARM** :

L'examen s'effectue yeux ouverts OU fermés. Les sujets sont assis dans une base de fauteuil roulant, chaque bras étant confortablement installé dans un bras d'exosquelette, avec des ajustements effectués pour s'adapter aux dimensions du corps du sujet. Un bras est détendu (bras passif) afin de permettre au robot de le déplacer dans 1 des 9 emplacements spatiaux différents. L'examineur demande ensuite au sujet de bouger son bras opposé (bras actif) afin de reproduire le mouvement du bras passif. Lorsque le sujet est au repos, le bras est détendu au niveau de la cible centrale (épaule en abduction horizontale à 30° et coude en flexion à 90°).

Voir également description *partie 1.1.4.3*.

7- Le **Wrist Position Sense Test** :

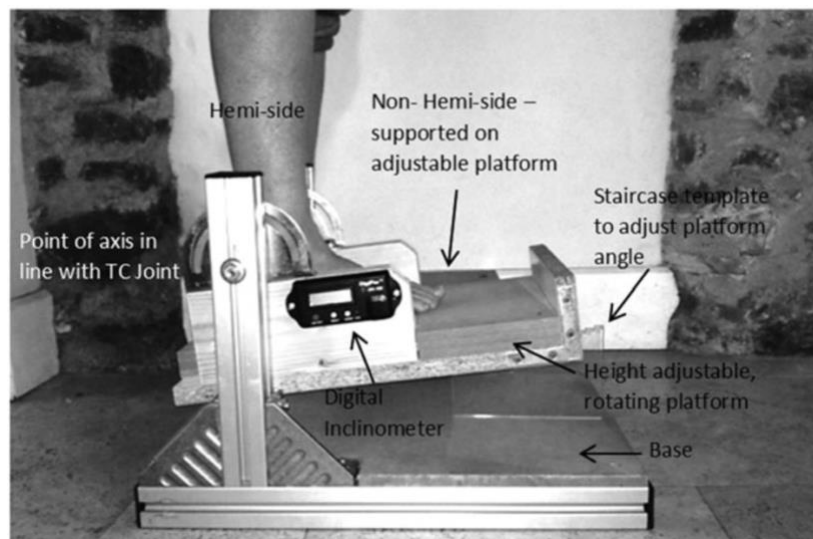


Le WPST quantifie la capacité à indiquer la position du poignet après un mouvement imposé. Les stimuli du test comprennent 20 angles prédéterminés du poignet en flexion-extension. Le dispositif comprend des échelles de rapporteur sur le couvercle et la base de l'appareil en forme de boîte, avec des marques d'échelle à intervalles réguliers, des attelles pour l'avant-bras et la main, ainsi qu'un rideau. L'attelle de l'avant-bras est fixée dans une position centrale alignant l'avant-bras et la main. L'attelle de la main est fixée à un levier, permettant un mouvement libre du poignet. Un pointeur, aligné sur l'axe du mouvement du poignet et fixé

au sommet de la boîte au-dessus d'un rapporteur, permet au sujet d'indiquer la position perçue du poignet. L'axe de mouvement du pointeur de réponse correspond exactement à celui du levier de test et les calibrages d'angle sont effectués en fonction des échelles du rapporteur. Un petit rouleau inséré dans la base du levier permet un mouvement en douceur et un tissu en feutre sous le levier élimine le bruit créé par les mouvements.

En suivant les instructions, l'examineur induit des mouvements du poignet dans des positions suivant une séquence aléatoire prédéfinie en utilisant le levier. Le poignet du sujet est soigneusement aligné avec l'axe de mouvement du levier. L'examineur tente de déplacer le levier à une vitesse constante. Les sujets indiquent l'angle perçu du poignet en alignant le pointeur de l'échelle supérieure du rapporteur avec une ligne imaginaire reliant le milieu du poignet à l'index. La vision du sujet sur la position de son poignet et sur les manipulations du levier par l'examineur est masquée par l'appareil. Un pré-test de la position du poignet permet de vérifier la compréhension des instructions. L'angle indiqué par le sujet est comparé à l'angle du levier pour déterminer l'erreur. Chaque test dure environ 5 minutes.

8- Le GradDT :



Ce nouveau test a été développé pour évaluer la perception de la position du pied et de la cheville lors de la mise en charge complète. L'appareil a été conçu pour être aussi simple que possible afin de minimiser les coûts de production et de maximiser la facilité d'administration. Il se compose d'une plate-forme rotative bidirectionnelle placée sous le pied testé pour produire une pente dans le plan sagittal. L'axe de rotation est aligné avec la malléole latérale qui correspond à l'axe biomécanique de rotation lorsque la cheville est en flexion dorsale et en flexion plantaire.

Le pied non testé est placé sur une plateforme horizontale statique, réglable en hauteur, reflétant la hauteur de la plate-forme inclinée lorsqu'elle est positionnée à 0° par rapport à l'horizontale. Le pied testé est positionné sur la plateforme rotative. Les participants doivent transférer leur poids sur le pied non testé, soulever le pied testé hors de la plate-forme et le

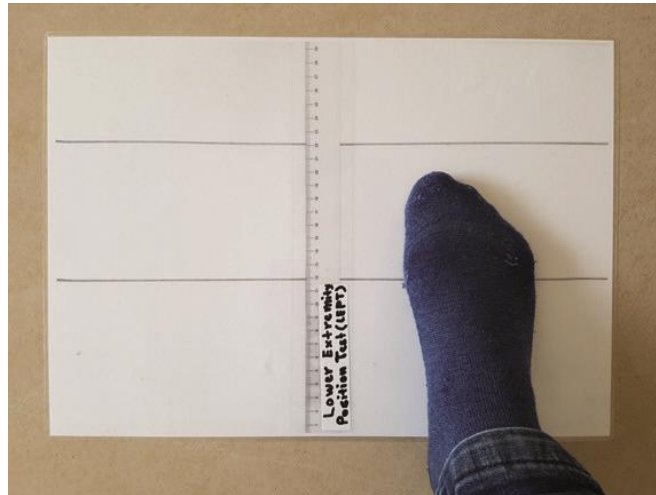
replacer une fois que la plate-forme avait été ajustée à la nouvelle position par l'évaluateur. Cela prend jusqu'à trois secondes. L'angle de la plateforme est ajusté manuellement, par l'évaluateur, à l'aide d'un gabarit d'escalier en acrylique découpé au laser, qui permet d'ajuster la pente de la plateforme rapidement, facilement et sans bruit. L'inclinaison et la précision de la pente sont confirmées par l'utilisation d'un inclinomètre numérique fixé à la plate-forme. La plateforme de base du stimulus est à 0° par rapport à l'horizontale et la pente de la surface de comparaison varie par incréments de 0,5° jusqu'à un maximum de 10° en flexion dorsale ou flexion plantaire. Le participant doit faire la distinction entre la plateforme de base et la plateforme de comparaison, en indiquant laquelle lui semble la plus inclinée. Les membres supérieurs sont soutenus pour faciliter l'équilibre et les participants doivent regarder droit devant eux et sans regarder leurs pieds.

9- Le **StepDT** :

Ce nouveau test a été conçu pour évaluer les seuils de discrimination du sens de la position des membres inférieurs. L'appareil est resté aussi simple que possible afin de minimiser les coûts de production et de maximiser la facilité d'administration. Le StepDT est composé d'une marche réglable, qui contient une série de marches de 6 mm s'emboîtant les unes dans les autres, permettant de régler la hauteur facilement, rapidement et silencieusement. La hauteur de la marche était de 100 mm (stimulus de base), ce qui correspond à la hauteur d'un trottoir standard, et la marche de comparaison est réglable par incréments de 6 mm jusqu'à une hauteur maximale de 154 mm. Le test implique que le participant transfère son poids sur le membre non testé tandis que l'évaluateur place passivement le membre testé sur la marche. Le participant devait faire la distinction entre la hauteur de la marche de base et celle de la marche de comparaison, en indiquant laquelle était la plus haute sous le membre testé. Le participant bénéficiait d'un soutien des membres supérieurs pour faciliter son équilibre et devait regarder droit devant lui pendant toute la durée du test pour éviter tout retour visuel de la position du membre.

10- Le **Lower Extremity Position Test** :

Le kit de test se compose d'une fine feuille de plastique (papier de format A3 en nylon) collée au sol. Une règle est placée au milieu pour mesurer la distance parcourue. Deux lignes sont tracées à 12 et 22 cm, perpendiculairement à la distance de la règle. Les distances de 12 et 22 cm ont été choisies sur la base du test de kinesthésie des membres supérieurs et après des essais avec des individus de taille et de longueur de jambes différentes.



Procédure de test : Le sujet a les yeux bandés et est assis sur une chaise, avec des chaussettes (pour éliminer autant que possible les frottements et les entrées tactiles). Une explication verbale et une démonstration visuelle utilisant le membre supérieur sont fournies pour s'assurer que les instructions soient comprises. Le test consiste en un lent mouvement passif du pied sur la surface en plastique, répété deux fois (la première pour sentir et percevoir et la seconde pour reproduire). Le mouvement est effectué par le testeur, à partir du bord de la surface en plastique (0 cm sur la règle), sans mouvement actif de la part du sujet, afin de permettre la réalisation du test pour les sujets atteints d'hémiplégie. Pour chaque répétition, le testeur déplace d'abord le pied de façon à ce que les orteils atteignent soit la ligne de 12 cm, soit la ligne de 22 cm, puis le ramène au point de départ ; enfin, il le déplace à nouveau jusqu'au point où le sujet dit "stop". Les ordres verbaux donnés aux participants sont les suivants "sens où va ton pied" lors du premier positionnement et "arrête-toi quand tu es arrivé au premier endroit" lors du mouvement de repositionnement. La différence entre la distance désignée et le point où le sujet demande d'arrêter le mouvement est mesurée en demi-centimètres (0,5 cm). Les sujets aphasiques peuvent utiliser un geste non verbal. Le test est répété deux fois pour chaque distance pour les deux pieds, dans un ordre aléatoire entre les distances et les pieds. Le décalage moyen est calculé pour chaque distance.

ANNEXE VII – Échelle AMSTAR 2

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

<p>1. Did the research questions and inclusion criteria for the review include the components of PICO?</p>		
<p>For Yes:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Population <input type="checkbox"/> Intervention <input type="checkbox"/> Comparator group <input type="checkbox"/> Outcome 	<p>Optional (recommended)</p> <ul style="list-style-type: none"> <input type="checkbox"/> Timeframe for follow-up 	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> No
<p>2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?</p>		
<p>For Partial Yes: The authors state that they had a written protocol or guide that included ALL the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> review question(s) <input type="checkbox"/> a search strategy <input type="checkbox"/> inclusion/exclusion criteria <input type="checkbox"/> a risk of bias assessment 	<p>For Yes: As for partial yes, plus the protocol should be registered and should also have specified:</p> <ul style="list-style-type: none"> <input type="checkbox"/> a meta-analysis/synthesis plan, if appropriate, <i>and</i> <input type="checkbox"/> a plan for investigating causes of heterogeneity <input type="checkbox"/> justification for any deviations from the protocol 	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No
<p>3. Did the review authors explain their selection of the study designs for inclusion in the review?</p>		
<p>For Yes, the review should satisfy ONE of the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> <i>Explanation for including only RCTs</i> <input type="checkbox"/> OR <i>Explanation for including only NRSI</i> <input type="checkbox"/> OR <i>Explanation for including both RCTs and NRSI</i> 		<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> No
<p>4. Did the review authors use a comprehensive literature search strategy?</p>		
<p>For Partial Yes (all the following):</p> <ul style="list-style-type: none"> <input type="checkbox"/> searched at least 2 databases (relevant to research question) <input type="checkbox"/> provided key word and/or search strategy <input type="checkbox"/> justified publication restrictions (e.g. language) 	<p>For Yes, should also have (all the following):</p> <ul style="list-style-type: none"> <input type="checkbox"/> searched the reference lists / bibliographies of included studies <input type="checkbox"/> searched trial/study registries <input type="checkbox"/> included/consulted content experts in the field <input type="checkbox"/> where relevant, searched for grey literature <input type="checkbox"/> conducted search within 24 months of completion of the review 	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No
<p>5. Did the review authors perform study selection in duplicate?</p>		
<p>For Yes, either ONE of the following:</p> <ul style="list-style-type: none"> <input type="checkbox"/> at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include <input type="checkbox"/> OR two reviewers selected a sample of eligible studies <i>and</i> achieved good agreement (at least 80 percent), with the remainder selected by one reviewer. 		<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

<p>6. Did the review authors perform data extraction in duplicate?</p>			
<p>For Yes, either ONE of the following:</p>			
<input type="checkbox"/> at least two reviewers achieved consensus on which data to extract from included studies		<input type="checkbox"/> Yes <input type="checkbox"/> No	
<input type="checkbox"/> OR two reviewers extracted data from a sample of eligible studies <u>and</u> achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.			
<p>7. Did the review authors provide a list of excluded studies and justify the exclusions?</p>			
<p>For Partial Yes:</p> <input type="checkbox"/> provided a list of all potentially relevant studies that were read in full-text form but excluded from the review	<p>For Yes, must also have:</p> <input type="checkbox"/> Justified the exclusion from the review of each potentially relevant study	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No	
<p>8. Did the review authors describe the included studies in adequate detail?</p>			
<p>For Partial Yes (ALL the following):</p> <input type="checkbox"/> described populations <input type="checkbox"/> described interventions <input type="checkbox"/> described comparators <input type="checkbox"/> described outcomes <input type="checkbox"/> described research designs	<p>For Yes, should also have ALL the following:</p> <input type="checkbox"/> described population in detail <input type="checkbox"/> described intervention in detail (including doses where relevant) <input type="checkbox"/> described comparator in detail (including doses where relevant) <input type="checkbox"/> described study's setting <input type="checkbox"/> timeframe for follow-up	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No	
<p>9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?</p>			
<p>RCTs</p>			
<p>For Partial Yes, must have assessed RoB from</p> <input type="checkbox"/> unconcealed allocation, <i>and</i> <input type="checkbox"/> lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)	<p>For Yes, must also have assessed RoB from:</p> <input type="checkbox"/> allocation sequence that was not truly random, <i>and</i> <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only NRSI	
<p>NRSI</p>			
<p>For Partial Yes, must have assessed RoB:</p> <input type="checkbox"/> from confounding, <i>and</i> <input type="checkbox"/> from selection bias	<p>For Yes, must also have assessed RoB:</p> <input type="checkbox"/> methods used to ascertain exposures and outcomes, <i>and</i> <input type="checkbox"/> selection of the reported result from among multiple measurements or analyses of a specified outcome	<input type="checkbox"/> Yes <input type="checkbox"/> Partial Yes <input type="checkbox"/> No <input type="checkbox"/> Includes only RCTs	
<p>10. Did the review authors report on the sources of funding for the studies included in the review?</p>			
<p>For Yes</p> <input type="checkbox"/> Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies			<input type="checkbox"/> Yes <input type="checkbox"/> No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

<p>11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?</p>	
<p>RCTs For Yes:</p>	
<p><input type="checkbox"/> The authors justified combining the data in a meta-analysis</p> <p><input type="checkbox"/> AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.</p> <p><input type="checkbox"/> AND investigated the causes of any heterogeneity</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> No meta-analysis conducted</p>
<p>For NRSI For Yes:</p>	
<p><input type="checkbox"/> The authors justified combining the data in a meta-analysis</p> <p><input type="checkbox"/> AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present</p> <p><input type="checkbox"/> AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available</p> <p><input type="checkbox"/> AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> No meta-analysis conducted</p>
<p>12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?</p>	
<p>For Yes:</p>	
<p><input type="checkbox"/> included only low risk of bias RCTs</p> <p><input type="checkbox"/> OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> No meta-analysis conducted</p>
<p>13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?</p>	
<p>For Yes:</p>	
<p><input type="checkbox"/> included only low risk of bias RCTs</p> <p><input type="checkbox"/> OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p>
<p>14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?</p>	
<p>For Yes:</p>	
<p><input type="checkbox"/> There was no significant heterogeneity in the results</p> <p><input type="checkbox"/> OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p>
<p>15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?</p>	
<p>For Yes:</p>	
<p><input type="checkbox"/> performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> No meta-analysis conducted</p>

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

- | | |
|---|------------------------------|
| <input type="checkbox"/> The authors reported no competing interests OR | <input type="checkbox"/> Yes |
| <input type="checkbox"/> The authors described their funding sources and how they managed potential conflicts of interest | <input type="checkbox"/> No |

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017 Sep 21;358:j4008.

RÉSUMÉ

CONTEXTE : L'AVC est une pathologie fréquente et en augmentation avec la longévité croissante. Il induit, outre des troubles moteurs et parmi les troubles sensitifs, des troubles de la sensibilité proprioceptive, qui permet de situer son corps dans l'espace. Il a été prouvé qu'une bonne rééducation motrice passe par une bonne rééducation proprioceptive, il est donc nécessaire d'évaluer correctement ces déficits lors des différents bilans. Ceci n'est possible qu'en présence d'outils fiables.

OBJECTIF : Faire un état des lieux de la littérature actuelle quant à la fiabilité des tests de proprioception chez le sujet post-AVC.

MÉTHODE : Les études portant sur les propriétés clinimétriques d'outils évaluant la proprioception sur des sujets-post AVC ont été recherchées. 8 bases de données ont été investiguées à l'aide d'équations approuvées par des chercheurs professionnels. La qualité méthodologie de ces études a été évaluée et l'extraction de données a été effectuée en double.

RÉSULTATS : Certains outils semblent présenter une très bonne fiabilité intra et inter-évaluateur comme la partie proprioception du Fugl-Meyer Assessment et le GradDT, parmi les 10 outils étudiés.

CONCLUSION : Il existe à ce jour peu d'études sur le sujet, celles existantes souffrent d'une qualité méthodologique médiocre, mais ces études sont plutôt récentes donc nous pouvons supposer une possible prise de conscience. Il serait pertinent de proposer plus d'études sur la fiabilité du Fugl-Meyer Assessment qui semblent se démarquer pour une utilisation clinique.

MOTS-CLÉS : PROPRIOCEPTION, AVC, ÉVALUATION, FIABILITE

ABSTRACT

BACKGROUND : Stroke is a frequent pathology and is increasing with increasing longevity. It induces, in addition to motor disorders and among the sensory disorders, disorders of proprioceptive sensitivity, which concerns the ability to locate one's body in space. It has been proven that a good motor rehabilitation requires a good proprioceptive rehabilitation, it is therefore necessary to correctly evaluate these deficits during the different assessments. This is only possible if reliable tools are available.

OBJECTIVE : Review the current literature about the reliability of proprioception tests for the post-stroke subject.

METHOD : Studies investigating the clinimetric properties of tools assessing proprioception in poststroke subjects were searched. 8 databases were investigated using equations approved by professional researchers. The methodological quality of these studies was assessed by the QAREL scale and duplicate data extraction was performed.

RESULTS : Some tools seem to obtain a very good intra and inter-rater reliability such as the Fugl-Meyer Assessment proprioception part and the KINARM, among the 10 tools studied.

CONCLUSION : To date, there are few studies on the subject, and those that do exist suffer from poor methodological quality, but these studies are rather recent, so we can assume that there is a possible awareness. It would be relevant to propose more studies on Fugl-Meyer Assessment, which seem to stand out for a clinical use.

KEYWORDS : PROPRIOCEPTION, STROKE, ASSESSMENT, RELIABILITY