



HAL
open science

Validité prédictive du National Institute of Health Stroke Scale (NIHSS) : revue systématique et méta-analyse

Sergio Pimenta Afonso

► To cite this version:

Sergio Pimenta Afonso. Validité prédictive du National Institute of Health Stroke Scale (NIHSS) : revue systématique et méta-analyse. Médecine humaine et pathologie. 2020. dumas-03153217

HAL Id: dumas-03153217

<https://dumas.ccsd.cnrs.fr/dumas-03153217>

Submitted on 26 Feb 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.

**Validité prédictive du National Institute of
Health Stroke Scale (NIHSS) : Revue
systématique et méta-analyse**

PIMENTA AFONSO Sergio

Directeur de mémoire : M. Stéphan ROSTAGNO

Je tiens tout d'abord à remercier l'ensemble des personnes qui m'ont permis de mener à bien mes études ainsi que ce mémoire. Une devise militaire illustre bien mon état d'esprit « seul on va plus vite, ensemble on va plus loin ».

Mr Stephan ROSTAGNO, Directeur de mémoire, kinésithérapeute et responsable adjoint à la clinique Saint Martin (Marseille), Enseignant à l'IFMK de Marseille, évaluateur méthodologique d'essais clinique PEDRO et co-responsable du groupe d'intérêt en neurologie, merci de m'avoir transmis la passion pour la recherche. Votre patience, vos connaissances ainsi que votre pédagogie m'ont permis de dépasser tous les problèmes rencontrés et d'aller à chaque fois plus loin.

Toute l'équipe pédagogique de l'école de formation en Masso-kinésithérapie de Marseille et les intervenants professionnels, merci de m'avoir transmis votre passion pour ce métier ainsi que toutes les connaissances nécessaires pour exercer mon métier. Merci à Mr Delafontaine pour ses conseils précieux.

L'équipe du service MPR de l'Hôpital d'Instruction des Armées Sainte Anne à Toulon, Marie Salmeron, Charline Gassot, Emeline Desaleux, Mathilde Charensol, Marie Françoise Silva merci pour votre soutien, pour tout ce que vous m'avez enseigné et pour votre amitié. J'ai beaucoup de chance de travailler à vos côtés.

L'ensemble des professionnels de l'hôpital Sainte Anne qui m'ont aidé, Mr Ducret (directeur des soins), Mme Hacault (cadre supérieur de santé), Professeur Faivre (chef de service de neurologie), Professeur Thefenne (chef de service MPR) les équipes de neurologie, radiologie et réanimation de l'hôpital, merci de m'avoir mis dans les meilleures conditions pour que je puisse étudier et réaliser mon mémoire.

L'ensemble des kinésithérapeutes que j'ai croisé durant ma formation, merci pour vos conseils et pour tout ce que vous m'avez appris.

Mes parents, vous êtes pour moi source d'inspiration, de motivation et de fierté. Vous m'avez toujours soutenu, toujours poussé à me dépasser et vous m'avez appris que rien n'était impossible si on s'en donne les moyens. Ce que je suis, je vous le dois. Je vous remercie pour tout et espère un jour être digne de vous.

Ma sœur, tu as toujours été présente et m'a apporté tout le soutien et la motivation dont j'avais besoin. Tu as toujours cru en moi et je te dois beaucoup. Ta réussite est pour moi source d'inspiration et j'ai toujours été très fier de la personne que tu es. Merci mon pitchoun.

Mon beau-frère, tu m'as toujours soutenu, encouragé et poussé à donner le meilleur de moi-même. Ton côté rigoureux, combatif et compétitif mais aussi ta joie et bonne humeur sont source de motivation et inspiration pour moi. Merci pour tout.

Ma famille, c'est auprès de vous que je puise ma force, ma sérénité et ma détermination. Dans les bons comme dans les mauvais moments je peux toujours compter sur vous. Vous m'êtes indispensable. Merci. Je remercie également Valérie pour son aide.

Ma femme, tu as été à mes côtés durant ces 4 ans. Tu as su être patiente, compréhensive mais aussi déterminée et une véritable locomotive dans les moments de doute. Merci à toi pour ton soutien, ton amour ainsi que ta force.

SOMMAIRE

1.	Introduction.....	1
1.1	Accident Vasculaire Cérébral.....	1
1.1.1	Épidémiologie.....	1
1.1.2	Physiopathologie.....	2
1.1.3	Etiologie.....	2
1.1.4	Facteurs de risque.....	3
1.1.5	Symptômes et classification.....	3
1.1.6	Prise en charge aigue.....	5
1.1.7	Prise en charge au long terme et séquelles.....	8
1.2	NIHSS.....	9
1.2.1	Origine.....	9
1.2.2	L'échelle validé par Société Français Neuro-Vasculaire (SFNV).....	10
1.2.3	Propriétés métrologiques.....	11
1.3	Modified Rankin Scale (mRS).....	12
1.3.1	Origine.....	12
1.3.2	L'échelle.....	12
1.3.3	Propriétés métrologiques.....	12
1.4	Objectifs de la revue de littérature et intérêt pour la profession de Masso-kinésithérapeute..	13
2.	Méthode.....	14
2.1	Critères d'éligibilité.....	14
2.1.1	Schéma d'étude.....	14
2.1.2	Participants.....	14
2.1.3	Facteur pronostique.....	14
2.1.4	Indicateur.....	14
2.2	Méthodologie de recherche.....	15
2.2.1	Sources documentaires.....	15
2.2.2	Equation de recherche.....	15
2.3	Extraction et analyse des données.....	16
2.3.1	Sélection des études.....	16
2.3.2	Extraction des données.....	16
2.3.3	Evaluation de la qualité méthodologique.....	17
2.3.4	Méthode de synthèse des résultats.....	18

3. Résultats	20
3.1 Description des études	20
3.1.1 Diagramme de flux	20
3.1.2 Etudes exclues	21
3.1.3 Etudes inclus.....	21
3.2 Risques de biais des études inclus.....	26
3.3 validité prédictive du NIHSS	28
4. Discussion	34
4.1 Analyse des principaux résultats	34
4.2 Applicabilité des résultats en pratique clinique	39
4.3 Qualité des preuves et limites de la revue	43
4.4 Biais potentiels de la revue	44
5. Conclusion	46
Bibliographie	48
Liste des figures.....	51
Lexique mathématique	52
Sommaire des annexes.....	55

1. Introduction

« Est-ce que je vais entièrement récupérer ? Est-ce que je vais pouvoir reprendre mon quotidien, jardiner, sortir mon chien, aller au marché ? est-ce que je serais dépendant des autres ? »

Tout soignant prenant en charge des patients ayant eu un AVC a déjà été confronté à ces questions de la part du patient ou de la famille. En tant que kinésithérapeute on se rend compte que c'est la principale préoccupation du patient et que l'on est souvent en difficulté pour y répondre.

Ces mêmes questions nous nous les posons car lors de la réalisation de notre bilan diagnostique kinésithérapique (BDK) et du choix du traitement, le potentiel de récupération fonctionnel du patient ainsi que son devenir seront fondamentaux à prendre en compte. Si le patient est potentiellement indépendant à domicile ou alité en structure spécialisée, notre prise en charge sera alors radicalement différente, spécifique et adaptée au patient. La personnalisation du traitement est importante en kinésithérapie et encore plus en neurologie. C'est afin de pouvoir répondre à ces questions que j'ai voulu étudier la validité prédictive de l'échelle NIHSS.

C'est l'échelle prédictive à la suite d'un AVC la plus utilisée (par les médecins et infirmières principalement) mais j'ai constaté dans mes différents lieux de stage que les kinésithérapeutes n'utilisent pas cet outil qui pourrait changer leur prise en charge des AVC. Ce dernier point m'a également intrigué et poussé à étudier ce score.

1.1 Accident Vasculaire Cérébral

Sources : Collège des enseignants de neurologie et INSERM.

L'accident vasculaire cérébral ou AVC est défini selon l'Organisation Mondiale de la Santé (OMS) par : « Un Accident Vasculaire Cérébral résulte de l'interruption de la circulation sanguine dans le cerveau, en général quand un vaisseau sanguin éclate ou est bloqué par un caillot. L'apport en oxygène et en nutriments est stoppé, ce qui endommage les tissus cérébraux ».

Ce terme désigne donc un déficit neurologique focal dont l'étiologie est variable. Le mot « accident » souligne le côté brutal et évolutif de l'atteinte même si l'AVC est, dans une majorité de cas, le résultat de troubles vasculaires ou cardiaques antérieurs. Sous la même terminologie AVC, et des présentations cliniques parfois analogues, sont regroupées des maladies assez différentes.

Les Accidents Ischémiques Transitoires (AIT) qui correspondent à un déficit neurologique bref, réversible, avec une imagerie normale (IRM de perfusion et de diffusion normale) ne sont pas des AVC à proprement parler. Ils constituent une entité distincte.

1.1.1 Épidémiologie

Selon les dernières données de l'Institut National de la Santé Et de la Recherche Médicale (INSERM) :

- En France, on dénombre chaque année plus de 140 000 nouveaux cas d'AVC, soit un toutes les quatre minutes
- L'AVC représente la première cause de handicap physique acquis de l'adulte
- La deuxième cause de démence (après la maladie d'Alzheimer) et la deuxième cause de mortalité (première chez la femme) avec 20% des personnes qui décèdent dans l'année suivant l'AVC

- L'AVC peut survenir à tout âge : si l'âge moyen de survenue d'un AVC est de 74 ans, 25% des patients ont moins de 65 ans et 10% moins de 45 ans
- Ces dernières années le nombre d'AVC affectant des personnes jeunes a augmenté de manière significative.

L'étude de Santé publique France [1] nous montre la répartition des AVC sur la période 2008-2014 selon le sexe, l'âge, la région, l'étiologie ...

L'AVC est un réel problème de santé publique tant d'un point de vue médical (incidence, taux de mortalité, séquelles, perte d'autonomie de la personne âgée...) que d'un point de vue financier et politique (coût et durée du traitement, perte d'autonomie de la personne âgée et handicap du sujet jeune...)

1.1.2 Physiopathologie

Le cerveau humain nécessite un apport sanguin constant en oxygène et en glucose. La spécificité de cet organe est l'absence de réserve de ces 2 éléments. Ainsi en cas de chute brutal du débit artériel il y aura une souffrance du parenchyme cérébral. Les conséquences de cet « accident » dépendront de l'artère touchée ainsi que de la mise en place d'anastomoses de suppléance.

Lors d'un AVC il y aura dès le départ :

- Une zone centrale nécrosée que l'on ne pourra pas récupérer
- Une zone périphérique appelée la zone de pénombre. Les lésions y sont réversibles si un traitement est mis en place rapidement. C'est la zone ciblée lors des traitements aigus et elle sera visible par une IRM de diffusion.

Il y a 2 grandes catégories d'AVC :

- AVC ischémiques 80 %
- AVC hémorragiques 20 %

1.1.3 Etiologie

L'AVC est souvent de causes multiples toutefois 25% des AVC restent d'origine inconnue. Il est néanmoins indispensable de rechercher l'étiologie afin d'éviter la récurrence.

- Les AVC ischémiques (80%) peuvent provenir :
 - D'une plaque d'athérosclérose (environ 30 % des infarctus cérébraux)
 - D'une cardiopathie emboligène (fibrillation atriale, rétrécissement mitral...)
 - D'une atteinte dite « lacunaire » qui représente 20% des AVC ischémiques et se caractérise principalement par l'occlusion d'une artériole profonde de petite taille. La cause peut être une artériopathie ou l'hypertension artérielle
 - D'une dissection des artères cervico-encéphaliques. C'est la cause la plus fréquente des AVC du sujet jeune (20%) et peuvent être dû à un traumatisme, à une artère pathologique (maladie d'Ehlers-Danlos), à une hypertension artérielle chronique, à une surcharge pondérale ou peuvent être spontanés
 - D'une vasoconstriction réversible favorisée par certains médicaments ou toxiques
 - D'une artérite (Maladie de Horton)
 - Autres causes plus rares comme les hémopathies ou la drépanocytose.

- Les hémorragies intra-cérébrale peuvent avoir pour origine :
 - Hypertension Artérielle (HTA) chronique 50% des cas
 - Les angiopathies amyloïdes qui sont fréquentes chez le sujet âgé
 - Rupture d'une malformation artérioveineuse (représente 5 à 10% des cas mais touche principalement les sujets jeunes)
 - Les troubles de l'hémostase, acquis (anti-coagulant ou alcoolisme par exemple) ou congénitaux (hémopathie...)
 - Tumeurs cérébrales (5 à 10%)
 - Autres causes : comme les artérites cérébrales, les méningo-encéphalites herpétiques

A noter toutefois une spécificité des AVC hémorragiques. Ils constituent un phénomène dynamique puisque les atteintes neurologiques peuvent être observées dans les premiers 48 heures dû à l'œdème, à la pression péri-lésionnelle et à l'expansion de l'hématome (notamment en cas de trouble de la coagulation). La prédiction du devenir fonctionnel du patient devra tenir compte de cette période particulière.

1.1.4 Facteurs de risque

Plusieurs facteurs de risques ont été identifiés permettant ainsi aux soignants ainsi qu'aux responsables politiques de mettre en place des campagnes de prévention.

- HTA : c'est le premier facteur de risque des AVC (risque relatif de 4)
- Tabac : agit principalement sur l'athérosclérose (RR = 2)
- Hypercholestérolémie : risque d'AVC ischémique (RR = 1,5)
- Diabète : avec un RR = 1,5.
- Alcoolisme chronique : Au-delà de 3 verres par jour augmentation significative du risque d'AVC hémorragique
- Migraine : surtout chez les jeunes femmes dans les migraines avec aura et en association avec les autres facteurs de risque cardio-vasculaires
- Contraception orale : l'augmentation du RR a été démontrée pour la prise de fortes doses d'hormones. La pilule microdosée utilisée en France provoque une faible augmentation du RR.
- Obésité, syndrome métabolique, absence d'activité physique, syndrome d'apnées obstructives du sommeil : leur effet négatif a été clairement démontré sur les atteintes cardiaques ischémiques. Concernant le risque d'AVC les études sont discordantes.

Ces facteurs de risque ont une valeur collective très importante. A titre individuel, on peut n'avoir aucun facteur de risque et un AVC grave (par fibrillation auriculaire notamment).

1.1.5 Symptômes et classification

On doit évoquer un AVC devant le triptyque suivant :

- Déficit neurologique central caractérisé soit par :
 - Une perte de fonction (motricité, parole, sensibilité, vue...)
 - Une production excessive (douleurs, paresthésies, clonies...) qui est beaucoup plus rare
- Une atteinte focale correspondant à un territoire cérébral (anatomique ou vasculaire)
- Le coté brutal, rapide et évolutif de « l'accident »



Fig. 1 : campagne de l'INPES

Les différentes campagnes d'information nationale (figure 1) mettent l'accent sur les signes les plus visibles et facilement identifiables :

- La paralysie faciale unilatérale
- L'atteinte motrice d'un membre
- Et les troubles de l'élocution

A noter qu'il y a des formes classiques (hémiplégie, aphasie...) et des formes trompeuses plus difficiles à identifier (confusion, troubles du champ visuel...).

Plusieurs classifications existent pour l'AVC : selon la gravité (en fonction du score NIHSS que l'on verra ultérieurement), selon le mécanisme, la durée et l'évolution...

La plus communément utilisée est la classification selon la topographie (figure 2). Nous avons donc les AVC :

- Ischémiques avec :
 - Infarctus cérébraux de la circulation antérieure : carotidiens (artère ophtalmique, cérébrale antérieure, cérébrale moyenne)
 - Infarctus cérébraux de la circulation postérieure : vertébrobasilaires (artères vertébrales, cérébelleuses et cérébrales postérieures)
 - Petits infarctus profonds (ou « lacunaire »)
 - Infarctus cérébraux jonctionnels (entre deux territoires artériels).
- Hémorragiques avec :
 - Hématomes profonds
 - Hématomes superficiels
 - Hématomes sous-tentoriels

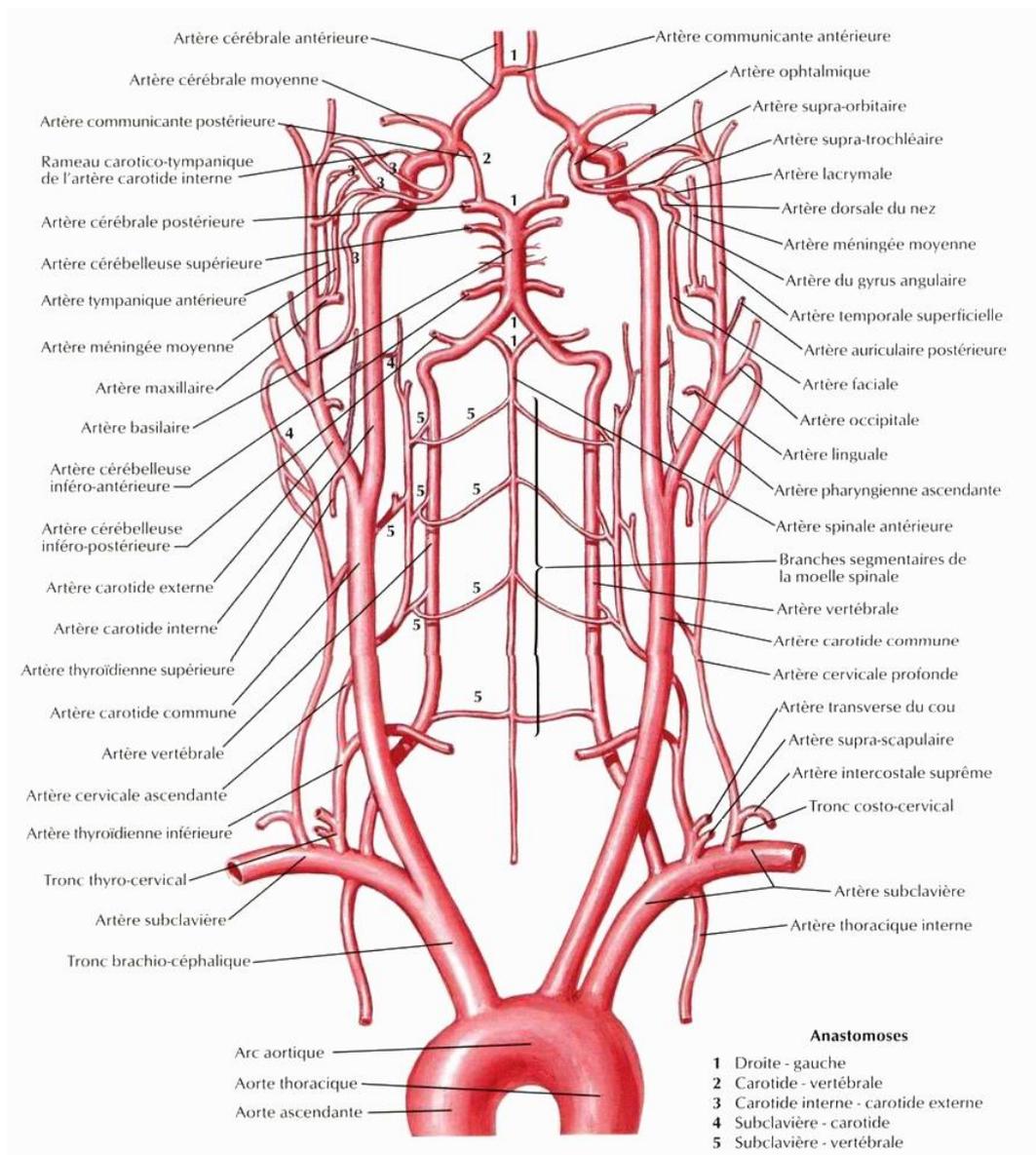


Fig. 2 : anatomie vasculaire cerveau (par Netter)

1.1.6 Prise en charge aiguë

La prise en charge d'un AVC est une urgence vitale absolue. La rapidité de prise en charge est un facteur primordial pour le devenir du patient.

Il y a trois temps importants lors de la prise en charge au sein des unités neurovasculaires :

- Le diagnostic de l'AVC
- Le traitement
- La recherche de l'étiologie pour éviter la récurrence

Le diagnostic :

- Il y aura tout d'abord un examen clinique objectivant les différents symptômes vus précédemment

- En cas de doute le premier examen sera un scanner (TDM) sans injection permettant de voir un saignement intra-cérébral puis un angioscanner cérébral (en temps artériel puis veineux) centré sur le polygone de Willis permettant de localiser la sténose artérielle voire veineuse. Le scanner cérébral est un bon outil pour visualiser les hémorragies cérébrales (figure 3). Il est moins sensible pour dépister les infarctus cérébraux précoces ou les infarctus de petites tailles.



Fig. 3 : Hématome intra-parenchymateux pariétal droit envahissant le ventricule droit

- L'examen de référence est l'IRM cérébrale (figures 4 et 5). C'est un examen diagnostique, pronostique et thérapeutique. Elle doit être réalisée dans les 4,5 heures afin que, si nécessaire, le patient éligible puisse bénéficier de la thrombolyse intra-veineuse. C'est un examen d'une durée de 20 minutes nécessitant que le patient soit immobile. Les premières séquences sont anatomiques : T1, T2 et T2*. Puis il y a la séquence de diffusion visualisant l'infarctus cérébral très précocement et la séquence de perfusion montrant la zone hypo-perfusée. La soustraction des images de ces 2 dernières séquences (mismatch) permettra de visualiser la zone de pénombre que l'on souhaitera revasculariser. Une dernière séquence TOF (Time of Flight) nous montrera l'arbre vasculaire individualisé. Le but de cette séquence est de visualiser la sténose mais aussi de permettre au médecin radiologue de prévoir le passage de son guide en cas de thrombectomie intra-artérielle. Toutefois cet examen est contre indiqué aux patients porteurs de pacemakers ou de neurostimulateurs par exemple et l'accès aux IRM peut être difficile.

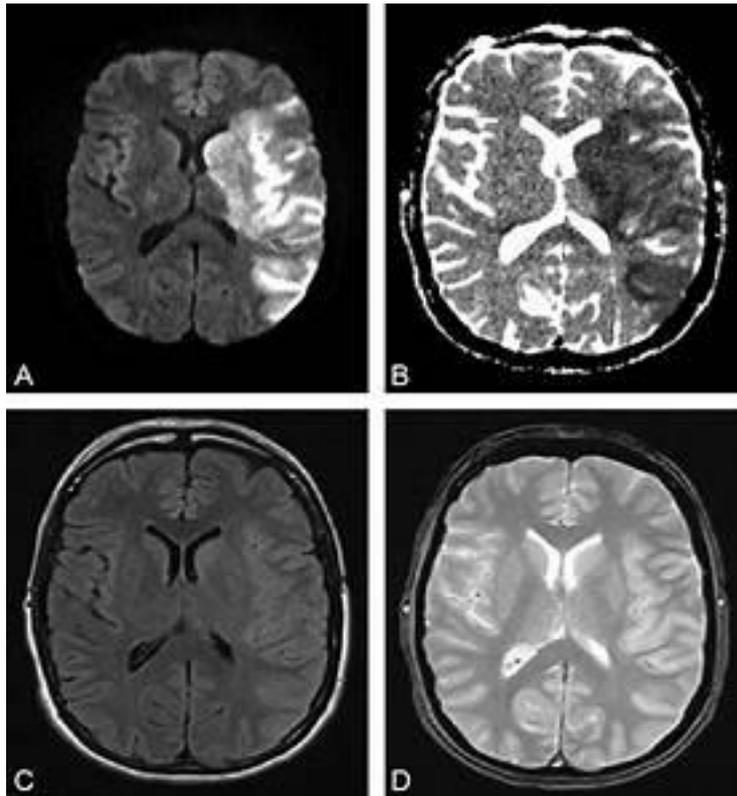


Fig. 4 : A-Diffusion ; B-perfusion ; C-T1 ; D-T2

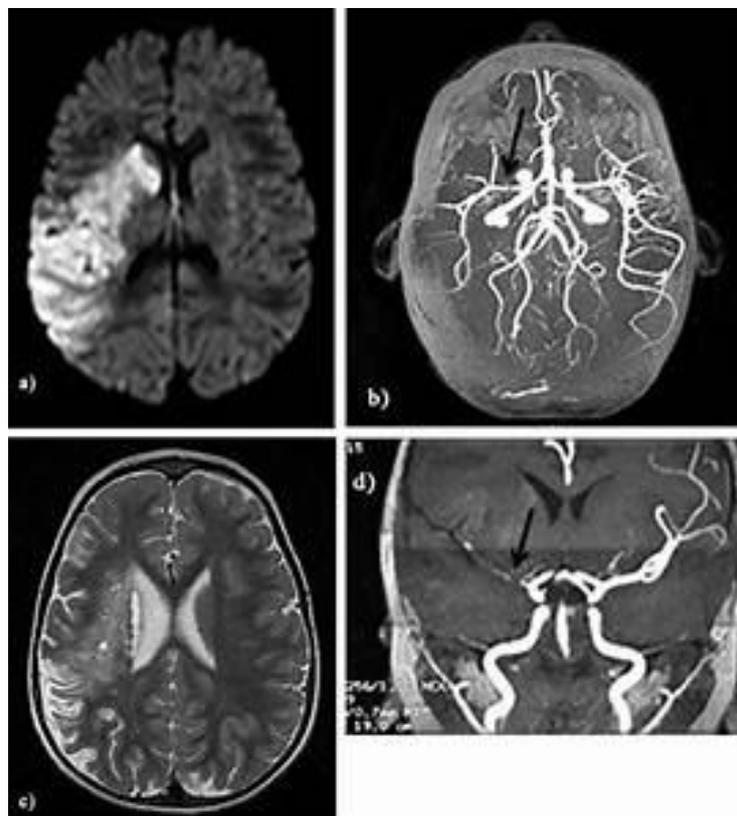


Fig. 5 : A-Diffusion ; B-TOF ; C-T2 ; D-TOF AVC ischémique sylvien droit

Le traitement

Il dépendra de l'examen clinique, du type d'AVC, du délai entre le début de l'AVC et la prise en charge, des différents examens ainsi que de la spécificité de l'hôpital (présence ou non de neuroradiologues spécialisés).

- Si la prise en charge est faite dans les 4,5 heures qui suivent les premiers signes le traitement pourra être une thrombolyse par le rt-PA (recombinant tissue-plasminogen activator) par voie intraveineuse. Il y a un risque d'hémorragie avec ce traitement, ainsi le neurologue responsable devra pondérer le rapport bénéfice/risque de cette injection.
- Dans les centres spécialisés le patient pourra bénéficier d'une thrombectomie intra-artérielle si les premiers signes remontent à moins de 6 heures.
- Plus rarement il y aura un acte de neurochirurgie en cas d'hématome intra-cérébral (afin de diminuer la pression intra-cérébrale et de diminuer les risques d'hydrocéphalie et d'engagement) ou en cas d'infarctus cérébral malin du sujet jeune.

Dans tous les cas il y aura une hospitalisation du patient en unité neurovasculaire et une réévaluation régulière de ce dernier (contrôle de ces constantes vitales, NIHSS etc.)

Bilan étiologique dans un deuxième temps

Le neurologue réalise l'ensemble des examens nécessaires afin de trouver l'étiologie : ECG, consultation en cardiologie, en urologie, bilan sanguin etc.

1.1.7 Prise en charge au long terme et séquelles

L'AVC est une atteinte grave, la mortalité est de 20 % à 1 mois et de 40 % à 1 an. La mortalité précoce est plus élevée en cas d'hémorragie intra-cérébrale que d'infarctus cérébral, en raison de l'effet de masse.

Concernant les survivants, schématiquement, un tiers sera dépendants, un tiers gardera des séquelles tout en étant indépendant et le dernier tiers retrouvera son état antérieur.

Le patient sera donc suivi pour :

- Éviter les complications (troubles cognitifs, spasticité, syndrome parkinsonien, troubles du décubitus etc.)
- Prendre en charge ces séquelles (motrices, sensibles etc.)
- Éviter la récurrence

L'essentiel de la récupération se fait dans les 3 voire 6 premiers mois. A l'issue, l'amélioration fonctionnelle est possible et tient davantage à une meilleure adaptation au handicap résiduel (rôle essentiel du kinésithérapeute).

En phase aiguë la préoccupation première du patient et de sa famille sera son devenir fonctionnel. Il est très difficile de répondre à cette question, c'est pour cela que le score NIHSS a une place prépondérante lors de la prise en charge aiguë des patients ayant eu un AVC.

1.2 NIHSS

1.2.1 Origine

Le National Institutes of Health (Instituts Américains de la Santé) correspond à un ensemble d'institutions gouvernementales des Etats Unis gérant la recherche médicale et biomédicale. Il fut créé en 1887 sous le nom de Laboratory of Hygiene puis réorganisé à 2 reprises : au début du 20^{ème} siècle et après la seconde guerre mondiale pour devenir le NIH.

De nos jours il est constitué d'une vingtaine d'instituts spécialisés dont le National Institute of Neurological Disorders and Stroke (NINDS) fondé en 1950. Les NIH sont constitués de deux parties, l'une dite intra-murale menant et finançant les recherches propres de l'institut, l'autre extra-murale soutenant des projets de recherche extérieurs aux NIH au sein des universités américaines.

Leur budget de 28,6 milliards d'euros permet le financement de 50 000 « Grants » (financement de projets) menés par 325 000 chercheurs au sein de 3 000 universités.

C'est justement un de ces projets en 1989 qui a permis à : Thomas Brott, MD, Harold P. Adams Jr., MD, Charles P. Olinger, MD, John R. Marler, MD, William G. Barsan, MD, Jose Biller, MD, Judith Spilker, RN, Renee Holleran, RN, Robert Eberle, Vicki Hertzberg, PhD, Marvin Rorick, MD, Charles J. Moomaw, PhD, and Michael Walker, MD³ [2] de publier un article proposant une échelle de 15 items pour examiner un patient en phase aigüe d'un AVC (figure 6). L'étude portant sur 24 patients a été publiée dans la revue Stroke et elle estime la validité et la fiabilité de cette échelle. Ces items sont basés sur 3 échelles utilisés auparavant : le Toronto Stroke Scale, le Oxbury Initial Severity Scale et le Cincinnati Stroke Scale.

Il est intéressant de noter que dès le début de l'article les auteurs précisent que l'échelle peut être utilisée par un neurologue, une infirmière des urgences ou une infirmière de neurologie tout en conservant la même fiabilité. Ceci montre bien que la préoccupation première des auteurs a été de faire une échelle fiable, valide et utilisable par les différents soignants.

1.2.2 L'échelle validé par la Société Français Neuro-Vasculaire (SFNV)

Score NIHSS

Item	Intitulé	cotation	score
1a	vigilance	0 vigilance normale, réactions vives 1 trouble léger de la vigilance : obnubilation, éveil plus ou moins adapté aux stimulations environnantes 2 coma ; réactions adaptées aux stimulations nociceptives 3 coma grave : réponse stéréotypée ou aucune réponse motrice	
1b	orientation (mois, âge)	0 deux réponses exactes 1 une seule bonne réponse 2 pas de bonne réponse	
1c	commandes (ouverture des yeux, ouverture du poing)	0 deux ordres effectués 1 un seul ordre effectué 2 aucun ordre effectué	
2	oculomotricité	0 oculomotricité normale 1 ophtalmoplégie partielle ou déviation réductible du regard 2 ophtalmoplégie horizontale complète ou déviation forcée du regard	
3	champ visuel	0 champ visuel normal 1 quadranopsie latérale homonyme ou hémianopsie incomplète ou négligence visuelle unilatérale 2 hémianopsie latérale homonyme franche 3 cécité bilatérale ou coma (la=3)	
4	paralysie faciale	0 motricité faciale normale 1 asymétrie faciale modérée (paralysie faciale unilatérale incomplète) 2 paralysie faciale unilatérale centrale franche 3 paralysie faciale périphérique ou diplégie faciale	
5	motricité membre supérieur	0 pas de déficit moteur proximal 1 affaissement dans les 10 secondes, mais sans atteindre le plan du lit. 2 effort contre la pesanteur, mais le membre chute dans les 10 secondes sur le plan du lit. 3 pas d'effort contre la pesanteur (le membre chute mais le patient peut faire un mouvement tel qu'une flexion de hanche ou une adduction.) 4 absence de mouvement (coter 4 si le patient ne fait aucun mouvement volontaire) X cotation impossible (amputation, arthrodèse)	Dt G
6	motricité membre inférieur	0 pas de déficit moteur proximal 1 affaissement dans les 5 secondes, mais sans atteindre le plan du lit. 2 effort contre la pesanteur, mais le membre chute dans les 5 secondes sur le plan du lit. 3 pas d'effort contre la pesanteur (le membre chute mais le patient peut faire un mouvement tel qu'une flexion de hanche ou une adduction.) 4 absence de mouvement (le patient ne fait aucun mouvement volontaire) X cotation impossible (amputation, arthrodèse)	Dt G
7	ataxie	0 ataxie absente 1 ataxie présente pour 1 membre 2 ataxie présente pour 2 membres ou plus	
8	sensibilité	0 sensibilité normale 1 hypoesthésie minime à modérée 2 hypoesthésie sévère ou anesthésie	
9	langage	0 pas d'aphasie 1 aphasie discrète à modérée : communication informative 2 aphasie sévère 3 mutisme ; aphasie totale	
10	dysarthrie	0 normal 1 dysarthrie discrète à modérée 2 dysarthrie sévère X cotation impossible	
11	extinction, négligence	0 absence d'extinction et de négligence 1 extinction dans une seule modalité, visuelle ou sensitive, ou négligence partielle auditive, spatiale ou personnelle. 2 négligence sévère ou anosognosie ou extinction portant sur plus d'une modalité sensorielle	
		TOTAL	

Fig. 6 : Score NIHSS

1.2.3 Propriétés métrologiques

Le NIHSS a été créé pour évaluer les patients à la suite d'un AVC en phase aiguë. Cette échelle cote le niveau de conscience, les mouvements extra oculaires, le champ de vision, le fonctionnement des muscles faciaux, la force des extrémités, le fonctionnement sensoriel, la coordination (ataxie), le langage (aphasie), le discours (dysarthrie) et l'héminégligence (négligence).

Le temps de réalisation de l'échelle est de moins de 10 minutes. Les scores totaux du NIHSS vont de 0 à 42, les valeurs plus hautes représentant des dommages cérébraux plus importants. Le niveau de sévérité de l'AVC est défini par les auteurs de l'article original selon les scores seuil de la manière suivante :

>20	AVC grave
15-20	AVC sévère
5-14	AVC modéré
< 5	AVC mineur

Fig. 7 : Classification des AVC selon la gravité

La fiabilité du test a été étudié à plusieurs reprises. Ainsi :

- Dans l'article original les auteurs retrouvent une fiabilité inter-examineur correcte (Kappa moyen = 0,69) et une fiabilité intra-examineur bonne (Kappa moyen = 0,77)
- Meyer (2002) [3] trouve, pour la fiabilité inter-examineur, une valeur du coefficient kappa excellente (kappa = 0,969). Il précise néanmoins que l'item concernant la dysarthrie est le moins fiable (Kappa = 0,289)
- Goldstein (1997) [4] va plus loin et évalue la fiabilité de cette échelle chez des non-neurologues. Il obtient une excellente fiabilité inter-évaluateur (coefficient de corrélation intra classe de 0,95) et une excellente fiabilité intra-examineur (ICC= 0,93)

Cette échelle présente donc une très bonne fiabilité inter et intra-examineur.

L'étude de cette fiabilité est indispensable avant de s'intéresser à la validité d'une échelle. En effet il est inutile de s'intéresser aux valeurs d'un outil de mesure si ce dernier n'a pas une bonne fiabilité inter et intra-examineur. Pour le NIHSS la fiabilité étant bonne nous allons pouvoir étudier sa validité.

La réactivité de cette échelle a peu été étudiée dans la littérature. Les effets de plancher et de plafond sont surtout importants à prendre en compte pour des scores diagnostiques. Ainsi ce paramètre ne sera pas développé dans ce travail.

Afin de maîtriser cette échelle une formation vidéo ainsi qu'une certification est possible sur <http://www.nihstrokescale.org/>.

Dans le cadre de ce mémoire je me suis formé à son utilisation grâce à des neurologues et réanimateurs certifiés.

L'explication détaillée de l'échelle est en annexe 1.

Les études que nous avons sélectionnées évaluent la validité prédictive du NIHSS en mettant en rapport le NIHSS initial et le score mRS à 3 ou 6 mois. Nous allons donc voir les propriétés de cette dernière échelle.

1.3 Modified Rankin Scale (mRS)

1.3.1 Origine

La Rankin Scale a été développé en Ecosse par Rankin [5] en 1975. Elle permet d'évaluer l'état fonctionnel du patient après un AVC en tenant compte de ces activités avant l'accident vasculaire cérébral. C'est une échelle rapide à mettre en place, simple et elle donne une image globale de l'état fonctionnel du patient.

Cette échelle a été modifié par la suite (lors d'une étude sur l'aspirine) en 1988 afin de rajouter le 0. Elle fut nommée Modified Rankin Scale (figure 8) par van Swieten, Koudstaal, Visser, Schouten et van Gijn (1988) [6].

1.3.2 L'échelle

Modified Rankin Scale (mRS)

Valeur	Symptomes
0	Aucun symptome
1	Pas d'incapacité en dehors des symptômes : activités et autonomie conservées
2	Handicap faible : incapable d'assurer les activités habituelles mais autonomie
3	Handicap modéré : besoin d'aide mais marche possible sans assistance
4	Handicap modérément sévère : marche et gestes quotidiens impossibles sans aide
5	Handicap majeur : alitement permanent, incontinence et soins de nursing permanent

Value	Symptoms
0	no symptoms at all
1	no significant disability, despite symptoms ; able to carry out all usual duties and activities
2	slight disability ; unable to carry out all previous activities but able to look after own affairs
3	moderate disability ; requiring some help, but able to walk without assistance
4	moderately severe disability ; unable to walk without assistance and unable to attend to own bodily needs without assistance
5	severe disability ; bedridden, incontinent and requiring constant nursing care and attention

Fig. 8 : Score mRS

1.3.3 Propriétés métrologiques

C'est un score rapide à obtenir (5 minutes) et qui peut être réalisé par les différents soignants intervenants.

- Sa Fiabilité a été étudiée plusieurs fois :
 - Wolfe, Taub, Woodrow et Burney (1991) [7] ont étudié 50 patients ayant eu un AVC. 2 infirmières ont évalué le patient à 2 reprises avec un intervalle de 2 à 3 semaines. Les fiabilités inter et intra-examineurs sont excellentes (Kappa = 0,95).
 - Plus récemment Shinohara, Minematsu, Amano et Ohashi (2006) [8] ont examiné la fidélité inter-examineur de la mRS. Vingt évaluateurs (neurologues et infirmières) ont visionné les vidéos de 30 patients interrogés, et ont noté chaque patient. La fidélité inter-examineur a été calculée en utilisant le coefficient de corrélation intra classe (CCI). L'étude a révélé que la fidélité inter-examineurs était excellente (CCI = 0,95 pour les neurologues et CCI = 0,96 pour les infirmières).
- Concernant la validité de cette échelle, un bon nombre d'étude comparent la mRS à l'échelle de Barthel ou l'échelle de mesure de l'indépendance fonctionnelle (MIF). Ainsi en 2004 Kwon, Harzema, Duncan et Min-Lai [9] ont examiné la validité concurrente de la mRS, de l'Indice de Barthel (IB), et du score Moteur de la Mesure de l'Indépendance Fonctionnelle (M-MIF), en utilisant des coefficients de corrélation de Spearman. D'excellentes corrélations ont été observées entre la mRS et l'IB ($r = 0,89$), et entre la MRS et le M-MIF ($r = 0,89$).

En conclusion l'échelle mRS présente une excellente fiabilité et validité. C'est le score utilisé dans toutes les études comme critère principal d'évolutivité des AVC, notamment dans les essais thérapeutiques.

1.4 Objectifs de la revue de littérature et intérêt pour la profession de Masso-kinésithérapeute

Nous avons donc vu l'importance des conséquences de l'AVC :

- Sur la santé de l'individu et de sa famille (atteinte potentiellement grave, séquelles et complications nombreuses)
- Épidémiologique (véritable problème de santé publique)
- Économique (coût de prise en charge).

En tant que kinésithérapeute nous allons avoir un rôle primordial pour diminuer les séquelles de la maladie mais aussi pour améliorer le devenir fonctionnel du patient.

Le temps est un facteur important, il est donc nécessaire que la kinésithérapie soit mise en place au plus tôt et de façon optimale. Lors de la réalisation du bilan kinésithérapique le choix du traitement dépend des altérations retrouvées, mais aussi du nouvel état neurologique du patient et de ses nouvelles capacités (suite aux atteintes neurologiques irréversibles). Ainsi le devenir fonctionnel du patient est un point clé à connaître afin de mettre en place des objectifs rééducatifs réalistes et les plus adaptés possible au patient (c'est le projet rééducatif personnalisé). Ceci correspond à la tendance actuelle de la médecine qui veut qu'un traitement soit appliqué au patient en fonction de ses singularités (génétiques, individuels, environnemental, psychologique etc.).

L'échelle NIHSS à la suite d'un AVC apparaît donc comme un outil important pour le kinésithérapeute. Cette dernière ayant une bonne fiabilité nous allons étudier sa validité prédictive en prenant pour référence l'évaluation fonctionnelle du patient à 3 ou 6 mois grâce à l'échelle mRS.

2. Méthode

La revue de la littérature est une méthode permettant de synthétiser les différents acquis de la littérature scientifique sur un sujet. Avoir une réponse à une question clinique en se basant sur les données actuelles de la science tel est l'objectif de ce travail. La lecture voire la rédaction d'une revue de littérature permet d'inscrire sa pratique professionnelle dans une démarche d'Evidence Based Practic.

Nous ferons une synthèse des différents articles ayant évalué la validité prédictive d'un facteur prédictif très utilisé : le NIHSS.

2.1 Critères d'éligibilité

2.1.1 Schéma d'étude

La revue de la littérature porte sur la validité prédictive du NIHSS. Nous sommes donc sur une question pronostique. Pour ce type de question les schémas d'études les plus adéquats sont l'étude de cohorte ou l'étude cas témoin (en cas de pathologie rare).

Dans le cas présent, nous sommes face à une pathologie fréquente et nous incluons exclusivement les études de cohorte portant sur la validité prédictive du NIHSS.

2.1.2 Participants

Les sujets inclus sont les patients ayant subi un AVC. Nous incluons les différentes :

- Etiologies : AVC ischémiques ou hémorragiques
- Topographies : circulation antérieure, postérieure ...
- Ages (jeune et âgée)
- Sexes (homme et femme)
- Traitements : thrombolyse, surveillance clinique, chirurgie...
- Tout type de NIHSS initial (AVC mineur ou grave)

2.1.3 Facteur pronostique

Nous allons étudier la capacité de l'échelle NIHSS à prédire le devenir du patient ayant eu un AVC. En tant que kinésithérapeute notre intérêt principal est son devenir fonctionnel à 3 mois, 6 mois ou 1 an.

Nous excluons les études portant sur le NIHSS modifié ou sur l'étude exclusive de la validité prédictive du NIHSS en association avec un autre facteur pronostique (NIHSS + âge par exemple sans étude du NIHSS seul). Les études portant sur la validité des items du NIHSS un par un et ne donnant pas de résultats sur le NIHSS dans sa totalité seront également exclus.

2.1.4 Indicateur

Il y a plusieurs indicateurs possibles permettant d'objectiver la capacité prédictive du NIHSS : l'échelle de Barthel, une IRM, l'échelle de GOS, le taux de mortalité ...

Le mRs est une échelle fonctionnelle ayant une bonne fiabilité et validité. De plus elle est facilement réalisable et très répandue. Nous choisirons donc cette échelle comme indicateur objectif de l'état fonctionnel du patient.

2.2 Méthodologie de recherche

2.2.1 Sources documentaires

Pour réaliser la revue de littérature les bases de données électroniques que nous consulterons sont : Pubmed et Science direct. Ce sont 2 bases de données complémentaires et regroupant un très grand nombre d'articles.

Les bases de données plus spécifiques tel que PEDRO ne sont pas aussi exhaustives concernant les questions pronostiques.

Nous chercherons également dans la littérature grise :

- Kinésithérapie scientifique
- Kinésithérapie La Revue
- Les thèses médicales, mémoires de recherche
- Congrès
- Articles de journaux
- Livres

2.2.2 Equation de recherche

Afin de mener à bien notre recherche sur les bases de données nous utiliseront le modèle PICO(S) (figure 9) pour définir les mots clés.

Patient	Stroke (AVC)
Intervention	NIHSS
Comparaison	mRs
Outcome	Devenir fonctionnel (outcome, prognosis, prognostic, prediction)
Study Design	Etude de cohorte

Fig. 9 : Question PICO

L'équation de recherche est basée sur le modèle PICO et comprend les termes suivants :

<u>Stroke</u>
AND
<u>NIHSS OR National Institute of Health Stroke Scale</u>
AND
<u>Outcome OR prognosis OR prognostic OR prediction</u>
AND
<u>mRs</u>

2.3 Extraction et analyse des données

2.3.1 Sélection des études

Deux lecteurs (le deuxième étant un pharmacien spécialisé en industrie) appliqueront l'équation et les critères de recherche afin de sélectionner les articles à inclure.

La recherche sera faite de façon indépendante puis à l'issue les articles sélectionnés par les 2 lecteurs seront inclus. Les articles retenus par seulement un des lecteurs seront débattus et en cas de litige l'avis sera demandé à un troisième lecteur expert. Cette double recherche permet de diminuer les oublis ou les articles choisis par erreur.

Les doublons seront supprimés puis le premier critère de sélection sera le titre. Ne seront retenus que les articles portant sur l'évaluation du NIHSS et non pas ceux utilisant le NIHSS comme indicateur (de l'efficacité d'un traitement par exemple).

Le deuxième critère de sélection sera l'abstract. Nous ne garderons que les études évaluant la validité prédictive du NIHSS (seul et entier) par validité de critère. Nous excluons les études évaluant la validité concurrente du NIHSS (comparaison de plusieurs échelles prédictives par exemple) ainsi que les études portant sur des scores de prédiction clinique incluant le NIHSS et sur des échelles modifiées du NIHSS.

Le dernier critère de recherche sera la partie méthode de l'étude. Nous ne garderons que les études de cohorte ayant pour indicateur le mRS. Les autres indicateurs ou critères de devenir fonctionnels seront exclus.

Pour terminer l'ensemble du texte sera lu pour confirmer que l'article répond bien à l'ensemble de nos critères.

2.3.2 Extraction des données

Après avoir obtenu l'ensemble des études correspondantes à notre question et nos critères de recherche l'extraction des données permettra de synthétiser les différentes informations importantes.

Les données extraites seront :

- Le nombre de patient étudié, l'âge moyen et l'effectif par sexe
- Le NIHSS médian initial, le délai de réalisation du NIHSS et le délai de réévaluation du patient
- Le schéma d'étude (prospective ou rétrospective)
- Le type d'AVC étudié
- Les critères de devenir favorable du patient
- La sensibilité, la spécificité, la valeur prédictive positive ou négative
- L'AUC ROC Curve et le seuil décisionnel (Cutoff) du NIHSS (Cf. Lexique mathématique)

Nous calculerons grâce à ces données le rapport de vraisemblance positif et négatif s'il n'est pas donné (Cf. Lexique mathématique).

2.3.3 Evaluation de la qualité méthodologique

La grille de lecture de référence pour une question pronostique est le Quips tools [10].

Les auteurs ont développé une grille de lecture en 6 parties (figure 10) avec pour chacune d'entre elles des critères permettant de classifier les études en faible, moyen ou haut risque de biais (Cf Annexe 4).



Fig. 10 : Quips Tools

- Participants : L'échantillon de l'étude doit représenter la population étudiée. Les critères sont : nombre de patients suffisants, description de l'origine des patients de l'échantillon, description des caractéristiques de l'échantillon en début d'étude, description adéquate des méthodes de recrutement et d'échantillonnage, description adéquate de la période et du lieu de recrutement et description adéquate des critères d'inclusion et d'exclusion
- Suivi : Les données disponibles (participants non perdus de vue) doivent bien représenter l'échantillon de départ. Les critères sont : taux de réponse adéquat des participants à l'étude, description des tentatives de collecte des données des patients perdus de vue, les motifs de perdus de vue sont donnés, description adéquate des participants perdus de vue et il n'y a pas de différences importantes entre les sujets ayant fini l'étude et les sujets perdus de vue

- Mesure du facteur pronostique (FP) : Le facteur pronostique doit être mesuré de la même manière pour tous les participants. Les critères sont : une définition ou une description claire du FP est fournie, la méthode de mesure du FP est fiable et valide, les variables continues sont rapportées ou bien les valeurs seuils sont appropriées, la méthode et le contexte de mesure du FP sont les mêmes pour tous les participants de l'étude, le FP est disponible pour une proportion suffisante de l'échantillon étudié et des méthodes adéquates de remplacement des données manquante sont utilisés.
- Mesure du critère de jugement : Le critère de jugement doit être mesuré de la même manière pour tous les patients. Les critères sont : une définition claire du critère de jugement est fournie, la méthode de mesure du critère de jugement est fiable et valide et la méthode et le contexte de mesure du critère de jugement sont les mêmes pour tous les participants
- Variables de confusion : Les potentiels facteurs de confusion doivent être correctement pris en compte. Les critères sont : tous les facteurs de confusion importants sont mesurés, des définitions claires des facteurs de confusion sont données, la mesure de tous les facteurs de confusions est fiable et valide, la méthode et le contexte de mesure des facteurs de confusion sont les mêmes pour tous les participants, des méthodes adéquates de remplacement des données sont utilisées pour les facteurs de confusion non disponibles, le schéma d'étude tient compte des facteurs de confusion importants et les facteurs de confusion importants sont pris en compte dans l'analyse
- Analyses statistiques et rapport des résultats : Les analyses statistiques doivent être correctes et tous les critères de jugement rapportés. Les critères sont : la présentation des données est suffisante pour évaluer la stratégie d'analyse, la stratégie de construction du modèle d'analyse est appropriée et basée sur une trame ou un modèle conceptuel, les méthodes d'analyse statistiques sont adéquates pour le schéma d'étude utilisé et la présentation des résultats n'est pas sélective.

2.3.4 Méthode de synthèse des résultats

La synthèse des résultats sera qualitative et quantitative.

L'ensemble des études seront évaluées qualitativement. Les études ayant un risque élevé de biais ne seront pas incluses dans l'analyse quantitative afin de diminuer les sources d'erreur de nos résultats.

De plus seules les études donnant dans les résultats une sensibilité, une spécificité et un AUC (accompagnées de l'intervalle de confiance) du test NIHSS seront retenues pour une évaluation quantitative.

Une méta-analyse sera effectuée sur les résultats des études donnant une AUC du ROC Curve avec un intervalle de confiance. C'est un indicateur de la qualité intrinsèque d'un test prédictif dont l'explication et l'intérêt est développé dans le lexique mathématique.

Le Rapport de Vraisemblance (RV+ et RV-) sera calculé quand il n'est pas donné.

En effet cet indicateur est plus intéressant que la sensibilité et la spécificité car il n'est pas dépendant de la prévalence de la maladie dans la population [11], il caractérise donc mieux un test pronostique. Il est calculé à partir de la sensibilité et la spécificité du NIHSS pour une valeur seuil (déterminée à partir de la ROC Curve grâce à l'indice de Youden). Un exemple d'utilisation de cet outil mathématique

est donné par l'étude de Garrison (2015) [12]. Une méta-analyse du RV+ et RV- ne sera réalisé que si les données sont suffisantes pour calculer l'intervalle de confiance. A titre indicatif et afin de montrer une tendance générale, une méta-analyse de la sensibilité et la spécificité sera réalisée.

Attention toutefois la sensibilité, spécificité, RV+ et RV- sont dépendant du seuil déterminé ou calculé par l'auteur. Ce Cutoff est différent selon les études donc une estimation qualitative de l'homogénéité de la méta-analyse sera donnée.

Ainsi une analyse qualitative et quantitative sera réalisée sur la validité prédictive du NIHSS et on apportera certaines nuances sur l'utilisation et l'interprétation du NIHSS en fonction :

- De l'âge
- Du moment de réalisation du NIHSS, voire de l'évolution de ce score lors des premiers jours qui suivent l'AVC
- Du territoire concerné
- Du traitement mise en place
- Du sexe
- Du type d'AVC
- De la sévérité de l'AVC

3. Résultats

3.1 Description des études

3.1.1 Diagramme de flux

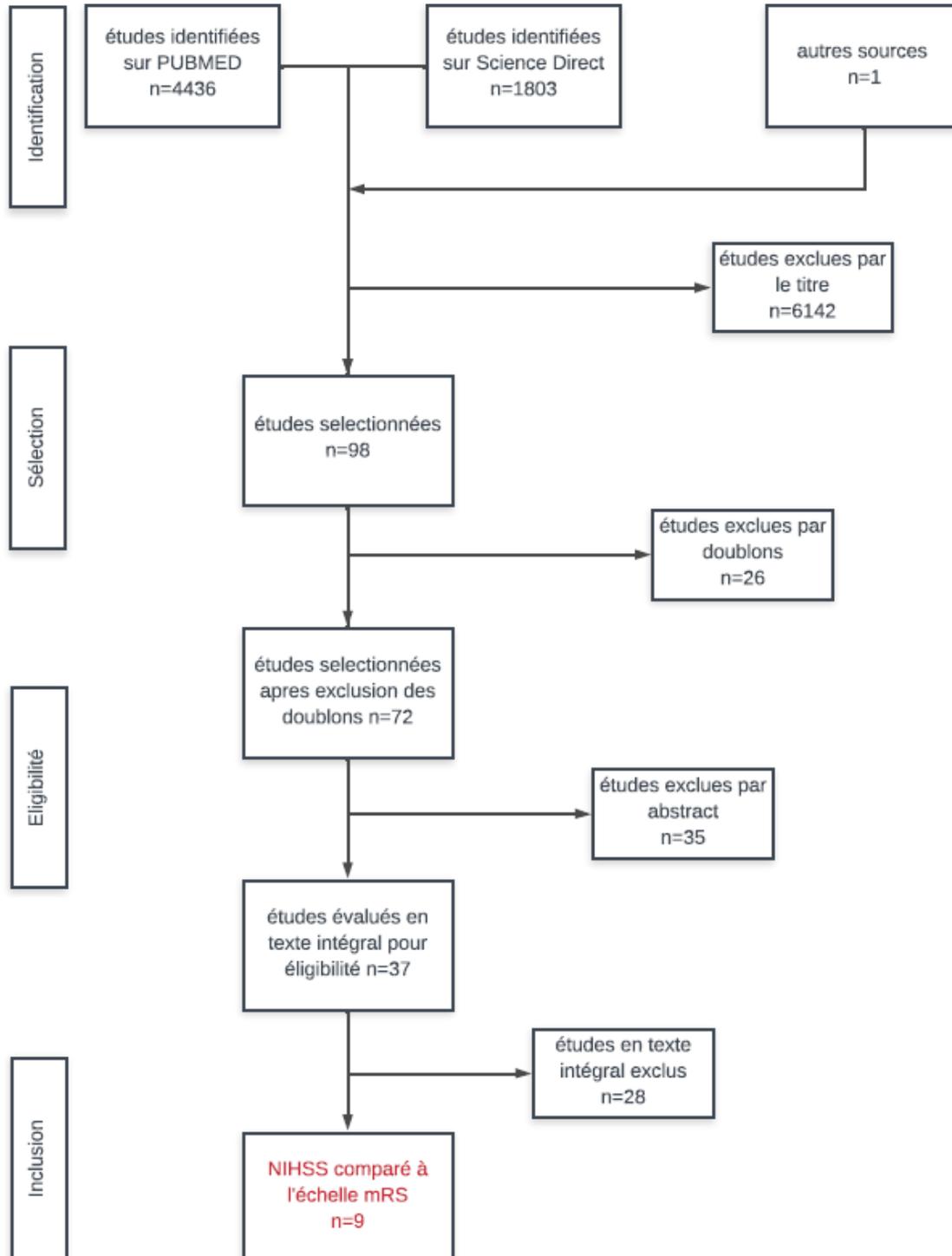


Fig. 11 : Diagramme de flux PRISMA

Grace à notre équation de recherche nous avons obtenu 4436 articles sur Pubmed et 1803 sur Science direct. Les autres sources nous ont permis d'ajouter un article supplémentaire. Le recueil des données fut réalisé entre le 1/06/2019 et le 1/11/2019.

Nous avons exclu 6142 études par le titre. La majorité des articles exclus utilisait le NIHSS comme indicateur de l'efficacité d'un traitement ou comme moyen de classification des patients avant application du traitement.

Sur les 98 études retenus 26 étaient en doublons ce qui nous donne 72 études.

35 études ont été exclus par l'abstract pour pouvoir obtenir 37 études évaluant la validité prédictive du NIHSS (seul et entier) par validité de critère.

Le dernier critère d'exclusion fut par la partie méthodologie des études. Seules les études ayant pour indicateur le mRS ont été retenus.

Le résultat final est : 9 études pour notre revue de la littérature (figure 11).

3.1.2 Etudes exclues

Notre méthodologie nous a amené à exclure 63 études.

Le tableau complet des études exclues est en annexe 2.

Tout d'abord 6142 études ont été exclues par le titre. Après avoir retiré les 26 études en doublon il restait 72 études.

En lisant l'abstract il fut possible de faire la classification suivante :

- 5 études portant sur la validité d'un score pronostique clinique incluant l'âge et le score NIHSS
- 7 études portant sur la validité du NIHSS modifié
- 3 études sur la fiabilité du NIHSS
- 2 études sur la prédiction de la dysphagie grâce au NIHSS
- 6 études comparant la validité prédictive du NIHSS et d'autres échelles
- 12 études comportant des données non exploitables (utilisation partielle du NIHSS, absence de résultat quantitatif)
- 37 études portant sur la validité par critère du NIHSS

Sur ses 37 études il a été possible de réaliser une dernière classification en fonction du critère de jugement (ou indicateur) utilisé :

- 2 études ayant pour indicateur l'échelle de Barthel
- 5 études ayant pour indicateur l'IRM
- 1 étude ayant pour indicateur l'échelle de GOS
- 20 études ayant pour indicateur d'autres critères (mortalité ou la capacité de la station bipodale du patient)
- **9 études ayant pour indicateur le mRS**

3.1.3 Etudes inclus

Le tableau complet est en annexe 3.

Voici le tableau simplifié avec les principales informations (figure 12).

Etudes /population	Effectif	Age Moyen	NIHSS initial médian	Délais réalisation NIHSS	Délais réévaluation /prédiction
Dubuc 2015 [13]	510	N.D.	NIHSS ≤3	N.D.	3 mois
Finocchi 2018 [14]	156	75.95 (13,1)	7 (4-16)	Admission	30 jours 3 mois
Mantero 2019 [15]	172	75.2 (13)	10 (4-10)	Admission	6 mois
Sato 2008 [16]	310	71	ant : 8(4-15) post : 4(2-7)	72h	3 mois
Sun 2011 [17]	49	67.41	NIHSS≤4 M=3	3h	mRS à la décharge
Wouters 2018 [18]	369 131	70.2 (57,6-78,1) 72.8 (61.5-79)	4 (2-7) 9 (7-15)	24h	3 mois
Wu 2019 [19]	269	62.68	6	Admission 2h ;24h ;7jours ; 10jours après traitement	3 mois
Yang 2014 [20]	1451	18-50 ans M=42.57	6	48h	mRS à la décharge
Yoshimura 2018 [21]	2902	67.80	N.D.	24h et 72h	3 mois

Fig. 12 : tableau récapitulatif simplifié de l'extraction des données des études

N.D. : Non Disponible

BI = Brain Infarction

ICH = IntraCerebral Haemorrhage

UNC= Undetermined type

T.I.V = Thrombolyse Intra-Veineuse

NIHSS≤4 ou régression rapide des symptômes

mRS à la décharge = Score mRS lors de la remise en charge du patient

L'extraction des données nous permet de voir les spécificités et singularités de chaque étude. Il est important de prendre en compte ces éléments lors de l'analyse des données quantitatives.

- ***Dubuc (2015)*** : Les auteurs se sont intéressés au score NIHSS chez les patients ayant eu un AIT ou un AVC ischémique mineur (NIHSS≤3). C'est une population intéressante car en dessous de 4 au NIHSS la thrombolyse intra-veineuse n'est pas administrée car le rapport bénéfice/risque n'est pas favorable. Les 510 patients ont été inclus de façon prospective au travers de l'étude CATCH. La question était de savoir si les score NIHSS à 90 jours était un bon indicateur du devenir fonctionnel du patient. Malheureusement les données mises à disposition sont insuffisantes pour avoir une vue globale des 510 patients inclus dans l'étude. Malgré les différents e-mails envoyés aux docteurs Dubuc, Choi et Hill, ainsi qu'à leurs départements de recherche respectifs nous n'avons pas reçu de réponse.

- Finocchi (2018) : Dans cette étude les auteurs se sont intéressés à la capacité prédictive du NIHSS chez les patients ayant eu un AVC hémorragique. En effet la validité prédictive du NIHSS pour les AVC ischémiques est communément accepté et bien documenté contrairement aux AVC hémorragiques. L'auteur a voulu démontrer dans son étude que le NIHSS était tout aussi valable pour prédire le devenir fonctionnel du patient pour les AVC hémorragiques. Cette étude de cohorte prospective inclus 156 patients (88 hommes et 68 femmes). L'âge moyen est de 75.95 ans et le NIHSS médian à l'admission est de 7. Le mRS est réalisé à 1 mois puis à 3 mois. Le critère de devenir fonctionnel favorable est un mRS \leq 3.

C'est l'étude ayant la moyenne d'âge la plus élevée. Elle porte exclusivement sur les patients ayant fait un AVC hémorragique.
- Mantero (2019) : Cette étude fait suite à celle de Finocchi (2018) réalisée l'année précédente. Les auteurs ont voulu étudier la validité prédictive du NIHSS à 6 mois pour un AVC hémorragique. Cette étude de cohorte prospective inclus 172 patients (94 hommes et 78 femmes). La moyenne d'âge est de 75.2 ans et le NIHSS médian à l'admission est de 10. Le mRS est réalisé à 6 mois et le critère de devenir fonctionnel favorable est mRS \leq 2.

L'étude porte donc également exclusivement sur les AVC hémorragiques mais cette fois-ci le délai d'évaluation du devenir fonctionnel est de 6 mois. A noter également que le score NIH à l'admission est plus élevé dans l'étude de Mantero (2019) que de Finocchi (2018).
- Sato (2008) : Dans cette étude les auteurs ont voulu comparer la validité prédictive du NIHSS pour un AVC ischémique antérieur et postérieur. En effet une étude précédente [22] montrait qu'un devenir fonctionnel favorable était plus commun chez les patients ayant eu un AVC postérieur. Ainsi les auteurs ont voulu comparer les capacités prédictives du NIHSS pour des AVC postérieurs et antérieurs et déterminer une valeur seuil pour un devenir fonctionnel favorable. Dans cette étude de cohorte prospective l'effectif est de 310 patients (202 hommes et 108 femmes). L'âge moyen est de 71 ans et le NIHSS, réalisé dans les 72 heures suivant l'admission du patient aux urgences, a une valeur médiane de 8 pour les AVC antérieurs et 4 pour les AVC postérieurs. Le devenir du patient à 3 mois est obtenu grâce au mRS et la valeur choisie pour un devenir fonctionnel favorable est mRS \leq 2. En plus d'évaluer la validité prédictive du NIHSS seul, les auteurs ont également pris en compte d'autres facteurs ayant une incidence sur le devenir du patient : âge, le sexe, l'hypertension artérielle, diabète, fibrillation auriculaire...

A noter que l'étude porte sur les AVC ischémiques n'ayant pas bénéficié de thrombolyse. Il faut également prendre en compte le fait que le score NIHSS est réalisé dans les 72 heures suivant l'admission du patient à l'hôpital et non pas immédiatement lors de l'admission.
- Sun (2011) : Les auteurs ont évalué les principaux facteurs prédictifs du devenir fonctionnel du patient ayant eu un AVC ischémique modéré ou avec une régression rapide des symptômes. Outre le NIHSS, d'autres facteurs sont évalués : âge, la fibrillation auriculaire, la sténose artérielle et la régression des symptômes. Seul les AVC ischémiques n'ayant pas bénéficié d'une thrombolyse y sont étudiés. Ceci est intéressant car dans de nombreuses études un des critères d'exclusion de la thrombolyse intra-veineuse est que le patient ai fait un AVC mineur et/ou ayant une régression symptomatique rapide. Or dans cette population un certain nombre de patient ont un devenir fonctionnel défavorable. Les auteurs ont donc voulu étudier la valeur prédictive du NIHSS chez ces patients pour ainsi mieux définir le seuil NIHSS pour une exclusion du traitement thrombolytique intra-veineux.

Cette étude de cohorte prospective inclus 49 patients (30 hommes et 19 femmes). L'âge moyen est de 67.41 ans et le NIHSS médian de 3. Le NIHSS réalisé dans les 3 heures qui suivent l'admission est inférieur ou égale à 4. Le devenir du patient est évalué lors de la remise en charge de ce dernier (verticalisation ou mise au fauteuil) et le critère de devenir fonctionnel favorable est un mRS ≤ 2 .

C'est l'étude portant sur l'échantillon le plus faible (49). La singularité de l'étude concerne sa population : seulement les AVC de sévérité mineure, modérés (NIHSS ≤ 4) ou ayant des symptômes régressant rapidement sont inclus. De plus le délai de réalisation du mRS n'est pas fixe puisque le neurologue le fera lors de la remise en charge du patient.

- Wouters (2018) : Cette étude est complètement différente des précédentes car elle s'intéresse à la validité prédictive du score NIHSS initial mais aussi à la validité prédictive et à l'intérêt de l'évolution de ce score les 24 premières heures. La question des auteurs est la suivante : est-ce que l'étude de la variation du NIHSS les 24 premières heures améliore la prédiction du devenir fonctionnel du patient ?

Cette étude de cohorte rétrospective inclut 369 patients (215 hommes et 154 femmes) ayant eu un AVC ischémique (bénéficiant ou non d'une thrombolyse intra-veineuse). L'âge moyen est de 70.2 ans. Le NIHSS initial appelé B-NIHSS, réalisé à l'admission a pour médiane 4. Le devenir fonctionnel est évalué à 3 mois grâce au mRS (le critère de devenir fonctionnel favorable est mRS ≤ 2). Ainsi les auteurs ont évalué la validité prédictive du B-NIHSS, du Delta-NIHSS (B-NIHSS moins NIHSS à 24h) et du RR-NIHSS (Delta-NIHSS divisé par B-NIHSS). D'autres indicateurs ont également été évalués : âge, sexe, pression artérielle, diabète, temps avant reperfusion, fumeur, fibrillation auriculaire.

La spécificité de cette étude est qu'en plus de l'étude de la validité prédictive du NIHSS sur l'ensemble de la population, les auteurs ont réalisé une analyse secondaire sur les patients ayant subi un AVC de gravité modéré ou sévère (NIHSS > 5). Ce sous-groupe est composé de 131 patients dont l'âge moyen est de 72.8 ans et le NIHSS initial médian est de 9.

Ainsi cette étude est complémentaire à l'étude précédente.

- Wu (2019) : Cette étude porte sur le devenir fonctionnel des patients à la suite d'un AVC ischémique ayant bénéficié d'une thrombolyse. Les auteurs souhaitent montrer que la validité du NIHSS est dépendante du temps. En effet la capacité du test à prédire le devenir du patient ne semble pas, intuitivement, la même que le NIHSS soit réalisé 2 heures après le début de l'AVC ou 10 jours après. Wu et Al. (2019) ont donc réalisé le NIHSS à l'admission, 2 heures, 24 heures, 7 jours et finalement 10 jours après le traitement. Ils ont donc évalué la validité prédictive de l'échelle à ces différents moments de réalisation et ont également calculé le seuil idéal (optimal Cutoff) pour chaque temps.

L'étude de cohorte prospective porte sur 269 patients (186 hommes et 83 femmes) dont la moyenne d'âge est de 62.68 ans. Le NIHSS médian à l'admission est de 6 mais il sera réévalué à 2 h, 24h, 7 jours et 10 jours post traitement. Le devenir fonctionnel favorable est défini par un mRS ≤ 3 à 3 mois. Les principaux attraits de cette étude sont qu'elle porte sur des patients ayant eu une thrombolyse puis qu'elle évalue la validité du NIHSS en fonction du moment de réalisation de ce test.

- Yang (2014) : Les auteurs de l'étude ont évalué la validité prédictive du NIHSS chez les patients jeunes (18-50 ans) ayant eu un AVC ischémique sans traitement par thrombolyse. Ils ont distingué les femmes des hommes afin de voir si la capacité à prédire le devenir fonctionnel des patients était différente en fonction du sexe pour cette population.

Ils ont également pris en compte les facteurs prédictifs secondaires (fibrillation auriculaire, type d'AVC, diabète et hypertension artérielle) afin de déterminer si leur fréquence est variable selon le sexe.

Ainsi cette étude de cohorte prospective inclus 1451 patients (995 hommes et 456 femmes). La moyenne d'âge est de 42.57 ans et le score NIHSS initial médian, réalisé dans les 48 heures qui suivent l'arrivée du patient, est de 6.

Tout comme Sun (2011) le mRS est évalué lors de la remise en charge du patient et le critère de devenir fonctionnel favorable est $mRS \leq 2$.

- Yoshimura (2018) : Cette étude a pour objectif de déterminer de façon distincte et comparative la validité prédictive pour les AVC infra et supratentoriels ainsi que de trouver le seuil NIHSS pour un devenir fonctionnel favorable.

Cette étude de cohorte rétrospective porte sur 2902 patients (1829 hommes et 1073 femmes). Ce schéma d'étude permet d'inclure plus de patients car il n'y a pas de délai d'attente entre l'AVC du patient et la réalisation du mRS. L'âge moyen est de 67.8 ans et le NIHSS initial moyen, réalisé dans les 72 premières heures, est de 8.8. L'indicateur réalisé à 3 mois est le mRS et le critère de devenir fonctionnel favorable est $mRS \leq 2$.

En conclusion chaque étude apporte une précision supplémentaire et des valeurs complémentaires sur la validité prédictive du NIHSS. Ainsi nous avons un spectre large et complet de patients et pathologies étudiés :

- AVC hémorragique/ischémique
- AVC ischémique au niveau de la circulation antérieure/postérieure
- AVC de gravité mineure à modérée/modérée à grave
- AVC avec thrombolyse I.V./AVC sans thrombolyse
- AVC chez les femmes/hommes
- AVC chez les patients jeunes/âgés
- AVC ischémiques supratentoriels/infratentoriels

3.2 Risques de biais des études incluses

Afin de pouvoir interpréter correctement les résultats il est fondamental de prendre en compte le risque de biais des études incluses (figures 13 et 14) grâce à la Quips tools (Cf. Annexe 4).

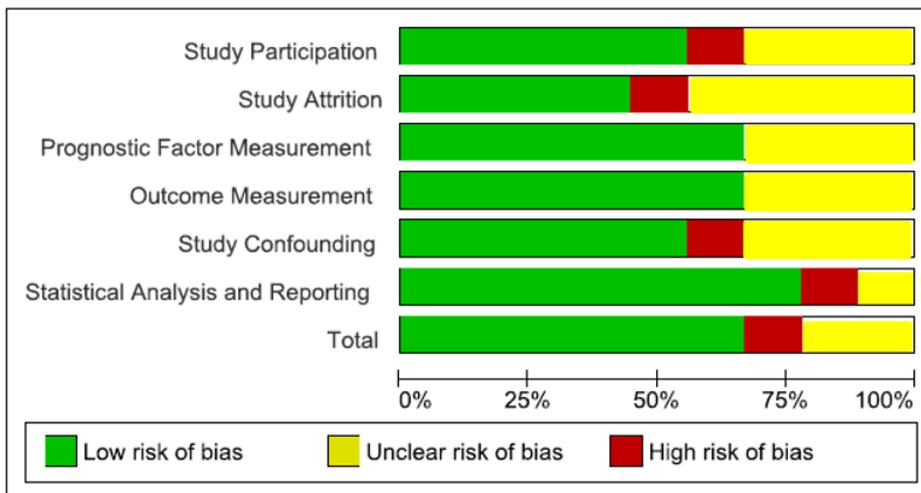


Fig. 13 : Schéma récapitulatif des biais

	Study Participation	Study Attrition	Prognostic Factor Measurement	Outcome Measurement	Study Confounding	Statistical Analysis and Reporting	Total
Dubuc 2015	●	●	?	?	●	●	●
Finocchi 2018	?	?	+	+	+	+	+
Mantero 2019	?	?	+	+	+	+	+
Sato 2008	+	?	+	+	?	+	+
Sun 2011	?	+	?	?	+	?	?
Wouters 2018	+	+	+	+	+	+	+
Wu 2019	+	+	+	+	+	+	+
Yang 2014	+	?	+	+	?	+	+
Yoshimura 2018	+	+	?	?	?	+	?

Fig. 14 : Biais des études incluses

- Dubuc (2015) : Cette étude présente un fort risque de biais. En effet au niveau des participants la description de l'origine des patients ainsi que des caractéristiques de l'échantillon en début d'étude ne sont pas donnés. La description de la méthode de recrutement ainsi que des critères d'inclusion et d'exclusion est également succincte. Au niveau du suivi, les auteurs ne font pas référence aux perdus de vue, il y a donc un fort risque que les données disponibles ne représentent pas bien l'échantillon de départ. Pour la mesure du facteur pronostique et du critère de jugement l'étude ne précise pas si les praticiens étaient formés aux 2 échelles et si l'application de ces dernières avait été identique pour tous les patients. La seule variable de confusion prise en compte est le traitement (exclusion des patients ayant eu une thrombolyse). Pour les analyses statistiques, l'étude ne rapporte qu'un tableau de concordance entre le NIHSS et le mRS à 90 jours. En résumé l'étude de Dubuc (2015) présente un fort risque de biais car les données mises à disposition sont insuffisantes. Malgré la demande des données brutes aux différents auteurs nous n'avons pas obtenu de réponse. Il faudra donc manipuler ces résultats avec précaution. Cette étude sera exclue de l'évaluation quantitative des résultats.
- Sun (2011) : Cette étude présente un risque de biais moyen. En effet le nombre de participant est faible (49) et il n'y a pas de description précise des critères d'inclusion et d'exclusion. Concernant la mesure du facteur pronostique et du critère de jugement l'auteur ne précise pas si la mesure a été identique pour tous les patients et si les évaluateurs ont été formés à ces échelles. Pour l'analyse statistique l'article nous donne l'Odd Ratio mais ne calcule pas la ROC Curve, AUC curve, le Cutoff, la sensibilité et spécificité.
- Yoshimura (2018) : Le risque de biais est moyen. Comme pour l'étude précédente, l'étude ne permet pas de dire si la mesure du facteur pronostique et du critère de jugement a été faite de manière identique pour tous les patients et si les évaluateurs ont été formés à l'utilisation des 2 échelles. De plus les auteurs ne tiennent pas compte du traitement des patients. C'est pourtant une variable de confusion majeure.
- Finocchi (2018) et Mantero (2019) : Ces 2 études sont très proches d'un point de vue méthodologie et présentent un risque de biais faible. Néanmoins la description des caractéristiques de l'échantillon en début d'étude est insuffisante et la description des participants perdus de vue est inexistante.
- Sato (2008) et Yang (2014) : Ces 2 études présentent un risque de biais faible. Il manque toutefois la description des motifs « perdus de vue » ainsi que des tentatives de collecte des données des patients perdus de vue. De plus les variables de confusion ne sont pas suffisamment développées.
- Wouters (2018) et Wu (2019) : Ces 2 études présentent les risques de biais les plus faibles. Selon le Quips Tools il n'y a pas de point majeur pouvant engendrer des biais.

En conclusion le biais principal des études que nous avons inclus est le biais d'attrition. En effet les perdus de vue sont parfois mal répertoriés voir mal décrit. Ce biais est important car il y a un risque que les données disponibles ne représentent pas bien l'échantillon de départ.

Ensuite il y a des biais possibles concernant les variables de confusion. La principale d'entre elles est le traitement. Ainsi le devenir du patient sera complètement différent en fonction du traitement dont il aura bénéficié. Il est donc important de prendre en compte le traitement des patients et de traiter séparément les patients ayant bénéficiés d'une thrombolyse de ceux n'ayant pas eu ce traitement.

Certaines études décrivent peu la population de départ ce qui peut être à l'origine de certains biais.

Les 2 échelles (NIHSS et mRS) ayant une bonne fiabilité, les risques de biais liés aux mesures des facteurs pronostiques et des critères de jugement sont faibles. Toutefois certaines études ne précisent pas si les évaluateurs sont formés à l'utilisation de ces échelles. Lyden et al. (1994) [26] ont montré que la formation à l'utilisation de l'échelle NIHSS améliorerait sa validité.

Les données statistiques et les rapports des résultats ne sont, quant à eux, que très peu à l'origine de risques de biais. A noter malgré tout qu'aucune étude ne donne directement les rapports de vraisemblance.

3.3 validité prédictive du NIHSS

Le tableau complet est en annexe 5.

Voici le tableau simplifié des principaux résultats (figure 15).

Etudes	RV +	RV-	AUC ROC Curve	Cut Off NIHSS
Dubuc 2015	N.D.	N.D.	N.D.	N.D.
Finocchi 2018	5,53	0,07	0,914	7
	3,03	0,11	0,833	7
Mantero 2019	3,40	0,17	0,886	10
Sato 2008	ant : 4,44 post : 4,42	ant : 0,24 post : 0,20	ant : 0,868 (0,818-0,917) post : 0,867 (0,783-0,951)	ant : 8 Post : 5
Sun 2011	N.D.	N.D.	N.D.	N.D.
Wouters 2018	2,45	0,28	0,78 (0,72-0,84)	7
	1,92	0,46		10
Wu 2019	1 : 3,19	1 : 0,58	1 : 0,76	1 : ≥ 12
	2 : 2,92	2 : 0,36	2 : 0,79	2 : ≥ 7
	3 : 2,37	3 : 0,26	3 : 0,83	3 : ≥ 5
	4 : 3,54	4 : 0,20	4 : 0,91	4 : ≥ 5
	5 : 4,25	5 : 0,19	5 : 0,92	5 : ≥ 5
Yang 2014	F : 3,71	F : 0,14	F : 0,881 (0,858-0,916)	F : 4
	H : 5,76	H : 0,24	H : 0,906 (0,886-0,923)	H : 6
Yoshimura 2018	S : 2,41	S : 0,48	S : 0,76 (0,74-0,78)	S : 10
	I : 1,64	I : 0,50	I : 0,69 (0,63-0,75)	I : 6

Fig. 15 : Tableau récapitulatif simplifié des résultats extraits

1= à l'admission

2= à 2 heures

3= à 24 heures

4= à 7 jours

5= à 10 jours

F= Femme

H= Homme

S= Supratentorial

I = Infratentorial

Concernant la validité prédictive du NIHSS il est important de constater que :

- Dubuc (2015) : Les auteurs sont partis du constat que parfois le score NIHSS est à 0 mais que le patient avait des déficits fonctionnels à 3 mois. Leur hypothèse est que le NIHSS n'est pas assez sensible pour les faibles déficits ainsi que pour les AVC de la circulation postérieure (Sato, 2008).
Nous n'avons accès qu'au tableau de contingence entre le NIHSS et le score mRS à 90 jours. De plus l'étude nous précise que la concordance complète (NIHSS et mRS à 0) est obtenue dans 54% des cas alors que la concordance partielle (NIHSS à 0 ou 1 et mRS à 0 ou 1) est obtenue dans 84% des cas. Nous n'avons pas le ROC Curve, de Cutoff, d'AUC, de sensibilité ou de spécificité. Cette étude ne sera donc pas retenue pour l'analyse quantitative.
- Finocchi (2018) : L'étude nous donne la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative, la ROC Curve et l'AUC. Grace à ces données nous avons calculé le RV+ et RV-. Il est important de noter que les seuils (Cutoff) NIHSS ont été choisis de façon subjective : les auteurs ont pris les valeurs médianes du NIHSS (7) et du mRS (4).
L'AUC est de 0.914 pour une prédiction à 1 mois et de 0.833 pour une prédiction à 3 mois.
Le RV+ est de 5.53 à 1 mois ce qui signifie qu'un patient avec un score NIHSS initial supérieur à 7 aura une probabilité 5,53 fois plus grande d'avoir un devenir fonctionnel défavorable à 1 mois. Le RV+ à 3 mois est de 3.03.
Le RV- à 1 mois est de 0.07. Ainsi un patient ayant un score NIHSS inférieur à 7 aura une probabilité 14 fois moins importante d'avoir un devenir fonctionnel défavorable à 1 mois. Pour une prédiction à 3 mois le RV- est de 0.11.
- Mantero (2019) : cette étude est la suite de celle vue précédemment. Elle nous donne en plus une validité du NIHSS pour une prédiction à 6 mois. Les outils mathématiques sont les mêmes à savoir : la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative, la ROC Curve et l'AUC. Les RV+ et RV- ont été calculés.
Ainsi pour une prédiction à 6 mois l'AUC est de 0.886. Le Cutoff est de 10. Le RV+ est de 3.4 et le RV- de 0.17.
- Sato (2008) : Les auteurs nous transmettent les mêmes données : la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative, la ROC Curve et l'AUC pour les AVC de la circulation antérieure et les AVC de la circulation postérieure. Nous avons calculé les RV+ et RV-.
L'AUC est très proche entre les AVC de la circulation antérieure et les AVC de la circulation postérieure (0.868 et 0.867 respectivement).
Il est en de même pour le RV+ avec 4.44 (antérieur) et 4.42 (postérieur) ainsi que pour le RV- avec, dans le même ordre, 0.24 et 0.20.
Le Cutoff Optimal a été cette fois-ci calculé grâce à la ROC Curve et l'indice de Youden. Il est de 5 pour un AVC postérieur et de 8 pour un AVC antérieur.
Ainsi le seuil pour un devenir fonctionnel favorable n'est pas le même pour un AVC postérieur et un AVC antérieur. Par exemple un score NIHSS de 7 prédit un devenir fonctionnel favorable pour un AVC de la circulation antérieure alors qu'il prédit un devenir fonctionnel défavorable pour un AVC de la circulation postérieure.
- Sun (2011) : Les auteurs nous fournissent un Odd Ratio avec intervalle de confiance à 95% des différents indicateurs prédictifs du devenir fonctionnel à la suite d'un AVC. Le score du NIHSS est de 1.44 (1.03-2.02). C'est l'OR le plus élevé n'incluant pas la valeur « 1 » dans l'intervalle

de confiance. La forme mathématique de ces résultats ne permet pas de les comparer aux autres études car à l'image de l'étude Dubuc (2015) les auteurs n'ont pas donné le ROC Curve, le Cutoff, l'AUC, la sensibilité ou la spécificité.

- Wouters (2018): Les principales données fournies par les auteurs sont : la sensibilité, spécificité, la ROC Curve et l'AUC. Les RV+ et RV- sont eux calculés.

Pour l'ensemble de la population l'AUC est de 0.78 pour le B-NIHSS. L'indice de Youden permet d'obtenir l'optimal Cutoff qui est de 7. Au niveau de ce seuil le RV+ est de 2.45 et le RV- est de 0.28.

Pour le sous-groupe composé exclusivement des patients ayant eu un AVC modéré ou sévère, le Cutoff est de 10. A ce seuil, le RV+ est de 1.92 et le RV- de 0.46.

A noter que les auteurs ont également comparé la validité prédictive du NIHSS chez des patients ayant bénéficié d'une thrombolyse intra-veineuse et ceux n'ayant pas eu ce traitement. Le B-NIHSS seul n'a pas été comparé mais un facteur complexe incluant le B-NIHSS, l'âge et le Delta-NIHSS nous est donné par l'auteur. Concernant ce facteur l'AUC (0.87 vs 0.86) est similaire dans les 2 populations.

L'étude nous montre une différence significative de la validité du NIHSS pour les patients ayant subi un AVC mineur et ceux ayant subi un AVC sévère.

- Wu (2019): Les données disponibles dans cette étude sont : la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative, la ROC Curve et l'AUC pour différents moments de réalisation du NIHSS. A posteriori nous avons calculé le RV+ et RV-.

Lorsque le NIHSS est réalisé à l'admission du patient, l'AUC est de 0.76. L'Optimal Cutoff obtenu grâce à l'indice de Youden est de 12. A ce niveau le RV+ est de 3.19 et le RV- de 0.58.

Lorsque le NIHSS est réalisé 2 heures après le traitement l'AUC est de 0.79. L'optimal Cutoff obtenu grâce à l'indice de Youden est de 7. A ce niveau le RV+ est de 2.92 et le RV- de 0.36.

Lorsque le NIHSS est réalisé 24 heures après le traitement l'AUC est de 0.83. L'optimal Cutoff obtenu grâce à l'indice de Youden est de 5. A ce niveau le RV+ est de 2.37 et le RV- de 0.26.

Lorsque le NIHSS est réalisé 7 jours après le traitement l'AUC est de 0.91. L'optimal Cutoff obtenu grâce à l'indice de Youden est de 5 également. A ce niveau le RV+ est de 3.54 et le RV- de 0.20.

Lorsque le NIHSS est réalisé 10 jours après le traitement l'AUC est de 0.92. L'optimal Cutoff obtenu grâce à l'indice de Youden est de 5 une nouvelle fois. A ce niveau le RV+ est de 4.25 et le RV- de 0.19.

Cette étude nous montre que la validité de l'échelle NIHSS est dépendante du temps.

- Yang (2014): Les auteurs nous donnent les données suivantes : la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative, la ROC Curve et l'AUC pour les jeunes femmes et les jeunes hommes ayant eu un AVC ischémique. Les RV+ et RV- ont été calculés grâce à ces données.

Pour les femmes, l'AUC est de 0.881 et le Cutoff obtenu grâce à l'indice de Youden est de 4. A ce seuil le RV+ est de 3.71 et le RV- est de 0.14.

Pour les hommes l'AUC est de 0.906 et le Cutoff optimal est de 6. Pour ce seuil le RV+ est de 5.76 et le RV- est de 0.24.

Cette étude nous montre une différence entre hommes et femmes. Par exemple un score NIHSS de 5 prédit un devenir fonctionnel favorable chez l'homme et défavorable chez la femme. De plus la validité de cette prédiction est meilleure chez l'homme que chez la femme.

- Yoshimura (2018) : Nous avons accès aux données suivantes : : la sensibilité, la spécificité, la ROC Curve et l'AUC pour les AVC supra et infratentoriels. Les auteurs ne calculent pas la valeur prédictive positive et négative. Ces données nous ont permis de calculer les RV+ et RV-. Concernant les AVC supratentoriels, L'AUC est de 0.76 et l'Optimal Cutoff calculé grâce à l'indice ce Youden est de 10. Pour ce score le RV+ est de 2.41 et le RV- de 0.48. Pour les AVC infratentoriels, l'AUC est de 0.69 et le Cutoff est de 6. Pour ce score NIHSS, le RV+ est de 1.64 et le RV- de 0.50. Cette étude nous montre bien une différence de validité prédictive du NIHSS pour un AVC supra et infratentoriel.

L'ensemble des données des études est insuffisant pour calculer l'intervalle de confiance du RV+ et RV-. Ainsi à titre indicatif on a réalisé une méta-analyse de la sensibilité (figure 16) et la spécificité (figure 17) du NIHSS. Il est important de noter que ces valeurs sont données dans les études en fonction du Cutoff choisi ou calculé. Les études sont donc hétérogènes pour réaliser la méta-analyse mais les résultats dégagent une tendance générale qui confirme les résultats décrit précédemment. Grace aux intervalles de confiance des sensibilités et spécificités, on confirme les tendances données par les RV+ et RV-.

Sensibilité NIHSS

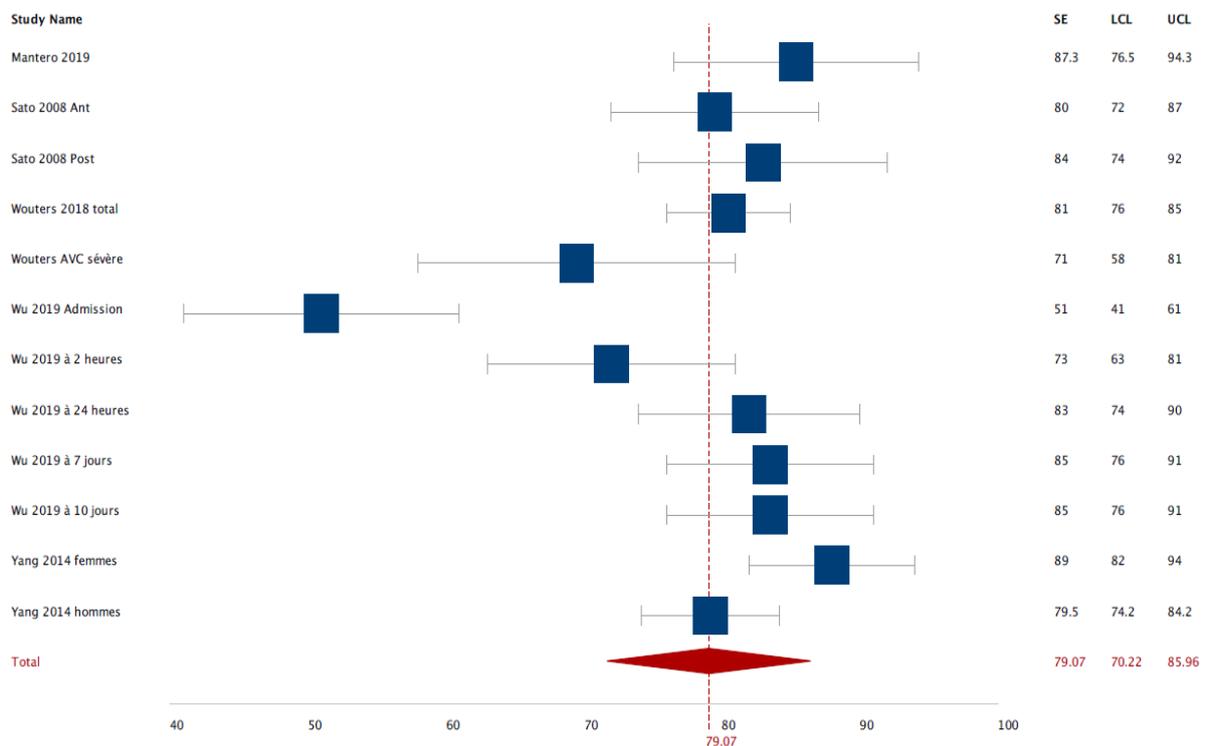


Fig. 16 : Méta-analyse de la sensibilité des études

Spécificité NIHSS

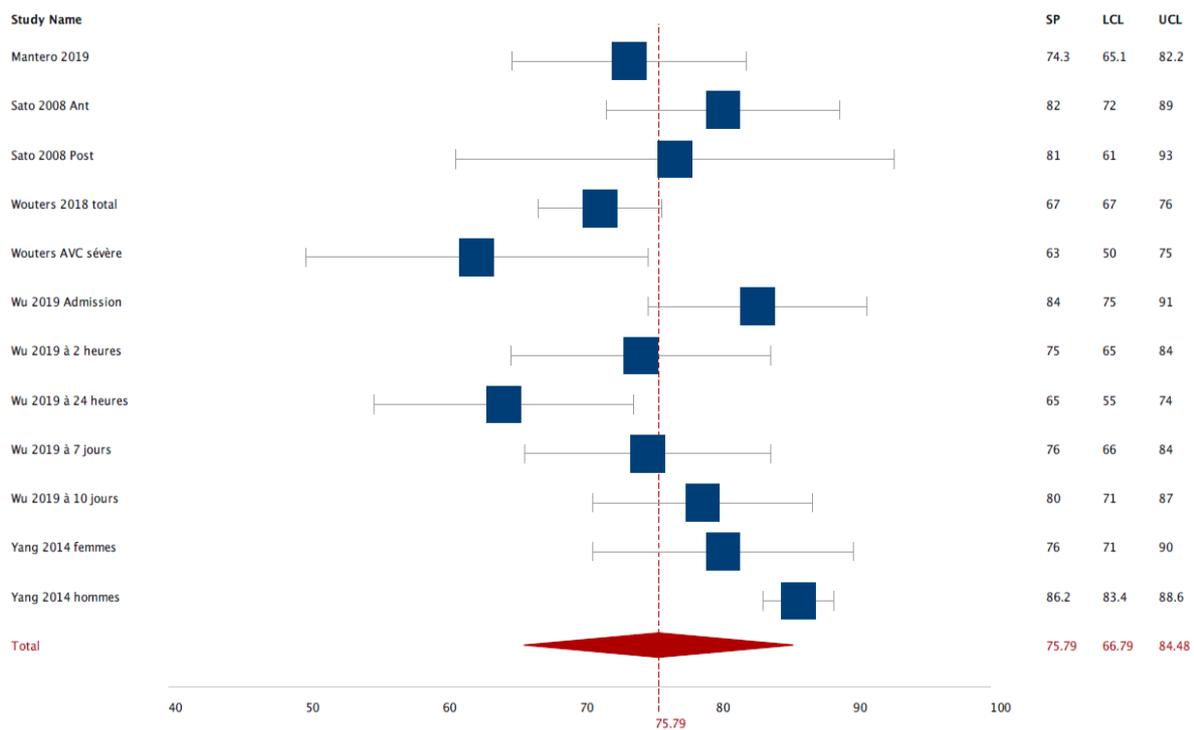


Fig. 17 : Méta-analyse de la spécificité des études

En plus de confirmer les résultats vus au-dessus ces 2 méta-analyses nous apportent des résultats complémentaires :

- Pour les AVC sévères, le NIHSS a une sensibilité et spécificité moindre
- Le NIHSS réalisé à l'admission et 2 heures après le traitement présente une meilleure spécificité que sensibilité
- Pour les femmes le NIHSS a une meilleure sensibilité que spécificité
- Pour les hommes c'est l'inverse, la spécificité est meilleure

Dans un deuxième temps on peut constater que Sato (2008), Wouters (2018), Yang (2014) et Yoshimura (2018) nous donnent l'AUC avec un intervalle de confiance. Nous pouvons donc faire une méta-analyse de cet indicateur global de la validité de l'échelle (figure 18).

AUC ROC Curve

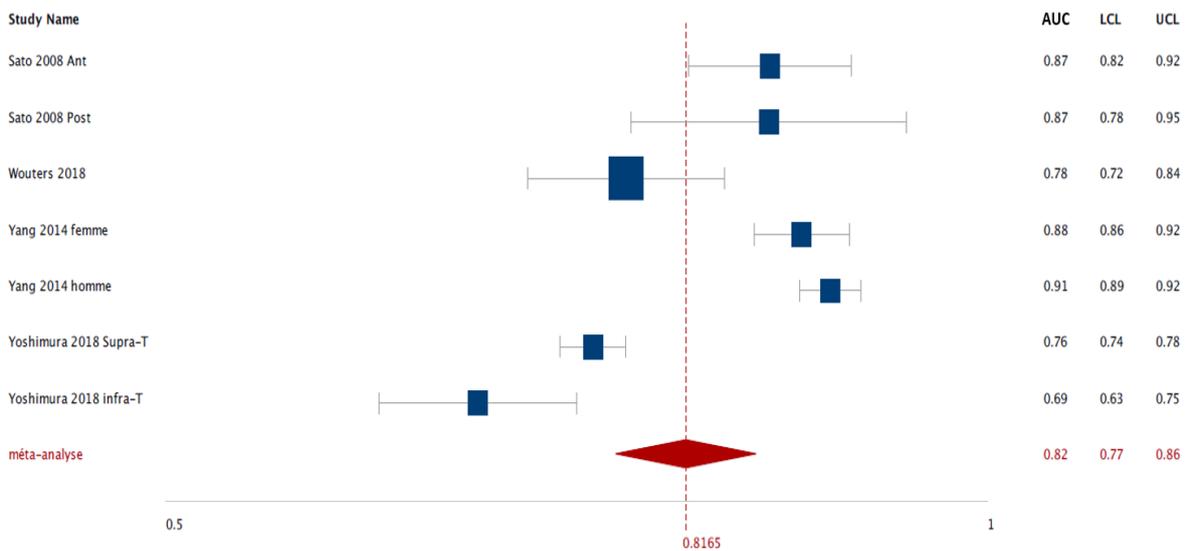


Fig. 18 : Méta-analyse de l'AUC des études

L'analyse des résultats de cette méta-analyse nous montre une AUC moyenne très élevée (0.8165).

L'AUC pour les patients ayant eu un AVC ischémique de la circulation antérieure (Sato, 2008 ant) ainsi que les patients jeunes (femmes et hommes Yang, 2014) est au-dessus de la moyenne.

A l'inverse, les patients ayant eu un AVC ischémique supra et infratentorial (Yoshimura, 2018) ainsi que les patients inclus dans l'étude Wouters (2018) se situent en dessous de la moyenne.

Il est important de noter que cette méta-analyse est hétérogène. En effet les études portent sur les populations différentes ainsi que des types d'AVC différents.

Liens d'intérêt

Je déclare ne pas avoir de liens d'intérêt pour la réalisation de cette revue de littérature.

4. Discussion

4.1 Analyse des principaux résultats

➤ La validité prédictive du NIHSS

La première observation que l'on peut faire est que le NIHSS a une bonne validité prédictive pour le devenir fonctionnel du patient à 1, 3 et 6 mois. Cette revue de littérature présentait l'intérêt de balayer un champ large de patients, de type d'AVC et de traitements mise en place. Ainsi le NIHSS à une bonne validité prédictive chez :

- AVC hémorragiques ou ischémiques
- AVC ischémiques au niveau de la circulation antérieure ou postérieure
- AVC de gravité mineure à modérée ou modérée à sévère
- AVC avec thrombolyse I.V ou AVC sans thrombolyse
- Patients jeunes ou âgés
- AVC ischémiques supratentoriels ou infratentoriels

L'AUC va de 0.69 à 0.92 ce qui montre bien que dans sa globalité cette échelle représente un excellent test. D'ailleurs l'AUC moyen est de 0.8165, le NIHSS a donc une excellente performance pronostique. Ainsi globalement pour les différentes valeurs du NIHSS la sensibilité et la spécificité sont très bonne.

Si on rentre plus dans le détail on constate que :

- La validité prédictive du NIHSS est similaire pour les AVC hémorragiques et les AVC ischémiques. C'est le grand apport des études de Finocchi (2018) et Mantero (2019). Le NIHSS a souvent été étudié pour les AVC ischémiques et son intérêt pour les AVC hémorragiques était moins clair. Ces 2 études nous montrent bien que les indicateurs de la validité prédictive (AUC, ROC Curve, sensibilité, spécificité ...) sont similaires pour ces 2 populations. Le NIHSS peut donc être aussi bien utilisé pour les AVC ischémiques que pour les AVC hémorragiques (une étude comparative AVC ischémiques/AVC hémorragiques est actuellement en cours de réalisation par Finocchi et son équipe). Finocchi (2018) va même plus loin et compare le NIHSS à la FUNC score et au ICH (Intracerebral hemorrhage) score (deux échelles très utilisées pour les AVC hémorragiques). Il nous soutient que le NIHSS est plus facile à réaliser, plus rapide et est purement clinique.
- La notion de temps est importante pour la validité prédictive du NIHSS. C'est une notion facile à comprendre intuitivement. Il semble plus facile de prédire ce qu'il va se passer dans un mois plutôt que dans 3 mois et j'ai moins de chance de me tromper si je réalise mon test 10 jours après l'AVC plutôt qu'à l'admission du patient, pour un devenir à 3 mois. Finocchi (2018) nous montre bien ceci puisque toutes les valeurs (AUC, RV+ et RV-) sont meilleures pour une prédiction à 1 mois plutôt qu'à 3 mois. Toutefois l'étude de Mantero (2019) contredit cette idée puisque les valeurs obtenues pour une prédiction à 6 mois sont légèrement plus élevées que celles présentées par Finocchi (2018) pour une prédiction à 3 mois. Cette incohérence peut s'expliquer par l'absence d'intervalle de confiance de ces valeurs. Nous ne pouvons pas affirmer avec certitude que la prédiction à 6 mois est meilleure que celle à 3 mois. De plus, comme le précise les auteurs, les populations ne sont pas identiques et dans l'étude de Mantero (2019) la médiane du NIHSS initial est substantiellement plus élevée.

De même, l'étude de Wu (2019) nous montre bien que la capacité du NIHSS à prédire le devenir fonctionnel du patient augmente avec le temps. En effet plus le NIHSS est réalisé à distance de l'AVC et meilleure est la prédiction à 3 mois (AUC, RV+ et RV- augmentent pour atteindre un pic à 10 jours). Fait très intéressant, la thrombolyse n'a pas une influence majeure sur ces valeurs. Elle modifie toutefois de façon importante le Cutoff (on passe de 12 avant la thrombolyse à 7 deux heures après le traitement). Il est important de relativiser ces résultats car le but de l'étude est de mieux utiliser le NIHSS pour mettre en place un traitement. Or certains traitements doivent être mis en place de façon urgente et ne permettent pas d'attendre 10 jours pour savoir le devenir fonctionnel du patient.

Les méta-analyses nous donnent également des pistes pour utiliser le NIHSS en fonction du temps. Réalisé à l'admission ou 2 heures après le traitement on va utiliser le NIHSS comme un examen d'inclusion (pour confirmer un devenir fonctionnel défavorable) alors qu'à partir de 24 heures après le traitement ce sera plutôt un examen d'exclusion (pour exclure un devenir fonctionnel défavorable). Cette notion est très importante en kinésithérapie. Un certain nombre de nos tests ou scores nous permettent d'exclure une pathologie (RV- proche de 0) et d'autres test nous permettent d'inclure un diagnostic (RV+ le plus grand possible). Dans une démarche EBP de diagnostique les premiers examens réalisés nous permettent d'exclure les pathologies les moins probable puis dans un deuxième temps il faudra des examens d'inclusion pour confirmer une pathologie parmi les plus probable.

- Un autre point important est la validité prédictive pour un AVC de la circulation postérieure. Plusieurs articles pointent du doigt les AVC postérieurs comme une limite du NIHSS [23]. En effet un devenir fonctionnel favorable est plus fréquemment rencontré chez les AVC postérieurs. C'est pour cette raison que Sato (2008) a réalisé son étude. On remarque que la validité prédictive est proche pour un AVC postérieur et un AVC antérieur. Le Cutoff lui est bien différents entre ces 2 populations.
- La sévérité initiale de l'AVC semble jouer un rôle important pour la validité prédictive du NIHSS. Dans l'étude de Wouters (2018) les différents indicateurs sont moins bons lorsque l'on prend seulement les patients ayant eu un AVC modéré ou sévère (NIHSS > 5). De même l'étude de Yoshimura (2018) a un score NIHSS initial moyen plus élevé (8.8) et présente des résultats (RV+, RV- et AUC) plus proches de ceux de Wouters (2018) et moins bons que les études portant sur les AVC ischémiques avec un score initial plus bas (Sato 2008, Wu 2019 et Yang 2014). Ceci est une première explication à la méta-analyse des AUC obtenus. En effet Yoshimura (2018) et Wouters (2018) sont en-dessous de la moyenne et Yang (2014) et Sato (2008) sont bien au-dessus. Le NIHSS a donc une moins bonne capacité prédictive lors d'AVC sévère.

Une explication possible de ce phénomène est que le taux de complications augmente avec la sévérité de l'AVC.

- Le sexe semble avoir un rôle sur la validité prédictive du NIHSS. Seule l'étude de Yang (2014) compare le NIHSS chez les femmes et chez les hommes. On constate un meilleur RV+ et AUC chez l'homme mais un meilleur RV- chez la femme. Donc la validité prédictive globale semble meilleure chez l'homme et l'utilisation du NIHSS sera différente chez la femme et chez l'homme. Le NIHSS pourrait être utilisé comme un examen d'inclusion chez l'homme (il confirmerait un devenir fonctionnel défavorable) alors que chez la femme ce serait un examen d'exclusion (il exclurait un devenir fonctionnel défavorable). Si l'on va plus loin dans l'analyse des résultats on peut dire que cet effet trouvé peut être simplement dû au hasard et qu'il n'est

pas possible d'affirmer avec certitude ces résultats. En effet nous ne pouvons pas calculer l'intervalle de confiance pour les RV+ et RV- mais nous avons celui de la sensibilité, la spécificité et l'AUC. Et pour ces valeurs la différence retrouvée entre hommes et femmes n'est pas significative puisque les valeurs pour ces 2 populations se chevauchent. Des études complémentaires et avec une population plus importante et plus large sont nécessaires afin de confirmer les résultats.

Toutefois la différence au niveau du Cutoff est significative. C'est sur ce dernier paramètre que nous pouvons affirmer qu'il y a une différence entre hommes et femmes ; on peut donc dire que pour un NIHSS bas la probabilité d'avoir un devenir fonctionnel défavorable est plus importante chez la femme. Ces résultats ne sont valables que chez les patients jeunes et des études complémentaires avec une population plus large sont nécessaires pour généraliser ces données.

Une ébauche d'explication de ces résultats est donnée par les auteurs. Les paralysies faciales, les dysarthries ainsi que les négligences sont plus souvent retrouvées chez les hommes. Ainsi le score NIH est plus élevé chez ces derniers.

- L'âge semble également influencer le NIHSS. Comme nous le montre le Forest Plot de l'AUC dans l'étude de Yang (2014) le NIHSS a une AUC plus élevée que pour les autres études. Ce paramètre est systématiquement utilisé lors de l'évaluation de scores de prédiction clinique mais à lui seul n'est pas un paramètre suffisant pour prédire le devenir fonctionnel du patient. En effet Sun (2011) trouve un Odd ratio pour l'âge de 1.03 (0.95-1.12). Ces résultats sont proches de ceux de Sato (2008) pour la circulation postérieure et antérieure (OR de 0.998 avec un intervalle de confiance de 0.959-1.038). De son côté, Wouters (2018) nous montre que l'AUC passe de 0.78 (0.72-0.84) à 0.82 (0.78-0.87) lorsque l'on inclut l'âge au score prédictif. La taille de l'effet n'est pas significative néanmoins ces études montrent une tendance qu'il sera important de confirmer par des études complémentaires.

Une explication possible est qu'avec l'âge, le nombre de facteurs de comorbidité augmente. Ainsi le nombre de complications est plus fréquents chez les personnes âgées en fonction du terrain.

- Pour finir la validité du NIHSS semble meilleure pour les AVC supratentoriels qu'infratentoriels. Ceci peut s'expliquer par l'échelle en elle-même. En effet il y a plus d'items évaluant la région supratentorielle (motricité, langage etc.) que la région infratentorielle. Ainsi le NIHSS sous-estime les AVC infratentoriels. De plus le Cutoff est plus bas pour cette population, on peut donc dire que pour un score NIHSS bas, la probabilité d'un devenir fonctionnel défavorable est plus importante chez les patients ayant fait un AVC infratentoriel.

Le NIHSS a donc une bonne validité prédictive mais il est important de connaître les différentes spécificités de cette échelle en fonction de la population afin d'interpréter et d'utiliser au mieux ce score.

A noter que la valeur seuil du mRS pour un devenir fonctionnel défavorable est subjective (2 ou 3 selon les auteurs). Nous pourrions nous interroger sur l'influence de ce seuil sur les résultats obtenus.

Or Sato (2008) a analysé la validité prédictive du NIHSS avec d'une part un seuil mRS à 3 et d'autres part un seuil à 2. Les valeurs obtenues sont similaires.

➤ Cutoff du NIHSS

Il est intuitif de comprendre qu'entre deux patients ayant bénéficié de ce test, celui ayant le score le plus bas aura un devenir fonctionnel plus favorable.

Mais si l'on ne compare pas deux patients, par exemple face au dossier d'un patient avec un score NIHSS de 7, que conclure ?

Le modèle idéal serait un modèle continu où chaque valeur du NIHSS serait associée à une probabilité de devenir fonctionnel favorable ou défavorable. Pour cela, il faut trouver une équation liant les 2 comme ce fut le cas dans l'étude de Joost J van Middendorp (2011) [24] pour les blessés médullaires. Dans cette étude, les auteurs ont trouvé une équation permettant de lier le score « Prediction Rule Score » à la probabilité de marche indépendante du patient. Ainsi pour chaque valeur du score une probabilité y est associée. L'obtention de cette équation n'est pas toujours possible et son applicabilité au quotidien difficile. C'est pour cette raison que déterminer un seuil pour une échelle quantitative continue est important.

Ce travail m'a permis de prendre conscience que lorsque l'on a une échelle quantitative continue, son utilisation implique de connaître les seuils significatifs (ou Optimal Cutoff) en fonction de la population, de la pathologie, du traitement ou du moment de réalisation de l'échelle. Par exemple le Time Up and Go test nous donne un temps de réalisation du test en seconde. Nous allons comparer ce temps aux temps de références/seuil afin de dire s'il y a un risque de chute ou non. Il en est de même pour le test de marche de 6 minutes par exemple. Ce seuil facilite notre pratique quotidienne. Cette solution de facilité a toutefois ses limites. En effet un patient se trouvant juste en dessous du seuil est très proche cliniquement d'un patient se trouvant juste au-dessus. Pourtant cet effet de seuil nous donnera des devenirs fonctionnels opposés.

- Sur l'ensemble des études, deux ne donnent pas de Cutoff : Dubuc (2015) et Sun (2011).
- Finocchi (2018) ainsi que Mantero (2019) ont choisi leur Cutoff de façon subjective car ils ont pris pour valeur seuil le score médian du NIHSS initial. Dans ces conditions on peut voir que le Cutoff est plus élevé pour une prédiction à 3 ou 6 mois (10) que pour une prédiction à 1 mois (7). Pour mieux comprendre, un patient ayant eu un AVC hémorragique avec un score NIHSS initial de 9 par exemple aura un devenir fonctionnel défavorable à 1 mois mais favorable à 3 ou 6 mois. Cette conclusion semble logique et intuitive puisqu'un AVC de plus grande gravité (NIHSS plus élevé) aura besoin de plus de temps pour récupérer. De plus la plasticité cérébrale, les traitements et la rééducation permettent une amélioration du patient entre 1 et 6 mois.
- Concernant les AVC de la circulation postérieure et antérieure on remarque que le Cutoff est plus élevé pour les AVC de la circulation antérieure (8 contre 5). Ainsi un score NIHSS de 7 par exemple nous indiquera un devenir fonctionnel favorable pour les AVC antérieurs et défavorable pour les AVC postérieurs. Comme Sato (2008) nous le précise « les AVC postérieurs ont une plus grande probabilité d'avoir un devenir fonctionnel défavorable à 3 mois pour un NIHSS relativement bas ». Cette différence peut s'expliquer par l'échelle NIHSS. L'hémisphère droit et l'hémisphère gauche sont évalués de manière différente par le NIHSS (en effet 7 points évalués ont un lien direct avec le langage alors que seulement 2 points sont en lien avec la négligence), tout comme les atteintes de la circulation antérieure et les atteintes de la circulation postérieure (plus d'items évaluent les fonctions motrices et corticales que les fonctions végétatives comme l'ataxie ou l'atteinte des nerfs crâniens). Ainsi le NIHSS sous-

estime les AVC postérieurs, il est important d'en tenir compte en prenant un seuil décisionnel plus bas.

- Chez les patients jeunes, le Cutoff est plus élevé chez les hommes que chez les femmes (6 vs 4). Un score de 5 par exemple prédit un devenir fonctionnel favorable chez l'homme mais un devenir fonctionnel défavorable chez la femme. Ainsi comme le précise l'auteur « les femmes ont une probabilité plus grande de devenir fonctionnel défavorable avec un score NIHSS plus bas ». L'auteur précise que les hommes ont plus fréquemment une paralysie faciale, une dysarthrie ou une négligence. Ceci fait que le score NIHSS chez les hommes sera automatiquement plus élevé. Il sera donc important de prendre en compte ce facteur lors de l'analyse du score NIHSS chez les hommes et femmes jeunes. Pour le moment aucune étude scientifique n'a généralisé ces résultats pour l'ensemble des hommes et femmes.
- La thrombolyse intra-veineuse est un événement majeur lors de la prise en charge d'un AVC ischémique. Que le NIHSS soit réalisé avant ou après ce traitement ne change pas significativement la validité prédictive du NIHSS. Toutefois le Cutoff est complètement différent (12 avant le traitement contre 7 deux heures après). Il sera essentiel de prendre ce facteur en compte. En effet un score NIHSS de 10 avant une thrombolyse sera prédictif d'un devenir fonctionnel favorable alors que ce même score après thrombolyse nous prédira l'inverse.
- Pour terminer le moment de réalisation du test a une importance pour la validité prédictive du NIHSS et pour la valeur du Cutoff Optimal. On voit que le Cutoff diminue entre l'admission du patient et 24 heures après le traitement par thrombolyse puis il atteint un plateau. La réalisation du NIHSS à ce moment semble être le plus intéressant. On voit ici l'effet de la re-canalisation puisque le Cutoff diminue entre l'admission et 2 heures après le traitement. Puis l'effet de la reperfusion car le seuil diminue encore jusqu'à 24 heures. Pour terminer au-delà de 24 heures il y a l'effet unité neuro-vasculaire. En effet de nombreuses études montrent que la prise en charge des AVC en unités spécifiques permet un meilleur gain fonctionnel pour le patient [38]. De plus l'étude de Wouters (2018) nous montre qu'inclure une variable supplémentaire (le delta-NIHSS) améliore la capacité prédictive du NIHSS. Le kinésithérapeute peut jouer un rôle à ce niveau. En effet le score NIHSS lors de l'admission aux urgences est réalisé par le neurologue, l'urgentiste voire le réanimateur. En revanche lors d'une prise en charge précoce du patient, le kinésithérapeute pourrait réaliser ce score à 24 heures pour ainsi avoir le delta-NIHSS.

On peut généraliser les conclusions de cette revue de la littérature à l'ensemble des échelles prédictives en médecine voire à l'ensemble des tests prédictifs et diagnostiques. En effet un score peut avoir une bonne fiabilité et une excellente validité mais il est important de savoir pour qui ? dans quel contexte ? comment ? pour quelles valeurs ? quel seuil ?

A noter que nous avons analysé le NIHSS en tant que variable unique pour prédire le devenir fonctionnel du patient. L'ensemble des études incluses sauf Wu (2019) ont également évaluées la validité prédictive du NIHSS associé à d'autres facteurs comme l'âge, la fibrillation auriculaire, le delta-NIHSS etc. (variable multiples). Ainsi il semblerait qu'en tenant compte d'autres facteurs, il soit possible d'améliorer la validité prédictive du NIHSS. Néanmoins le choix des facteurs ainsi que la confirmation de l'apport de ces facteurs sur la capacité à prédire du NIHSS est à prouver ultérieurement par une revue de la littérature. C'est ce qu'on appelle les règles ou scores de prédictions cliniques [25]. Dans

cet article les auteurs nous parlent de ces scores, de leurs intérêts en médecine ainsi que de leurs limites.

Dans la littérature de nombreuses études portent sur des échelles prédictives qui ont pour but d'aider à la mise en place d'un traitement. C'est le cas pour l'étude de J. J. van Middendorp (2011) [24] portant sur la validité prédictive d'un score pour la marche chez les patients traumatisés spinaux. Il y a un parallèle entre le NIHSS pour les AVC et le score ASIA pour les blessés médullaires. La méthodologie que nous avons développée est identique pour la validité prédictive du score ASIA pour les blessés médullaires.

- Les différents résultats vus précédemment sur la validité prédictive du NIHSS ainsi que sur le Cutoff sont spécifiques à chaque population étudiée et complexes. Afin de rendre plus simple l'utilisation et l'interprétation du NIHSS par le plus grand nombre nous proposons un schéma récapitulatif imprimable (figure 19). Ainsi en consultant ce schéma il est possible de visualiser rapidement la validité spécifique du NIHSS pour chaque population et type d'AVC ainsi que le Cutoff correspondant.

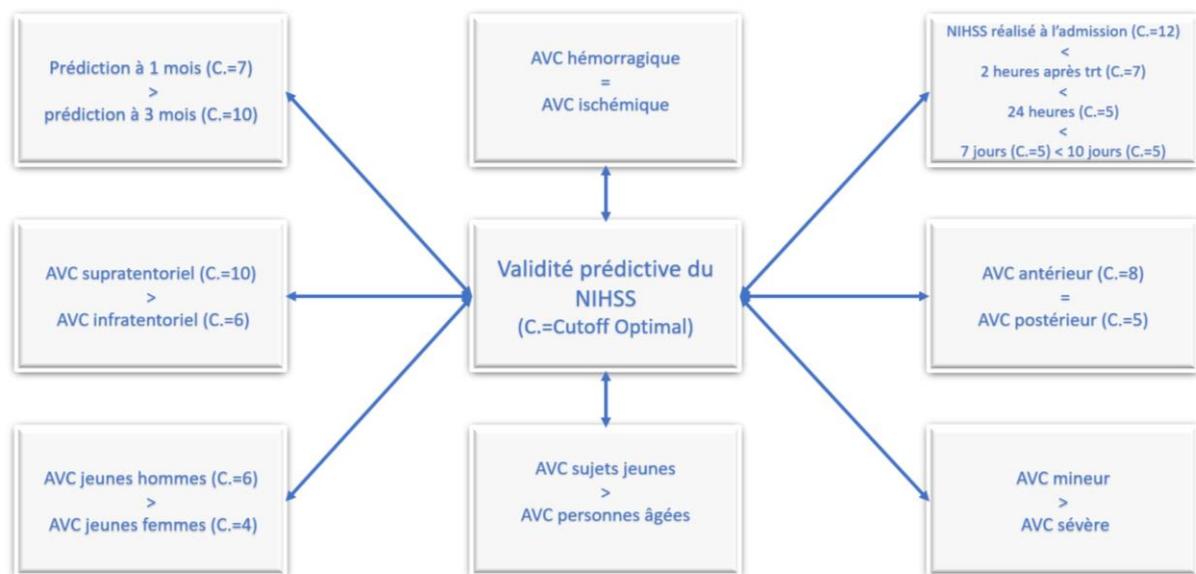


Fig. 19 : schéma récapitulatif

4.2 Applicabilité des résultats en pratique clinique

- Un score très largement utilisé

Dans le cadre de notre revue de littérature nous avons réalisé un sondage (Cf. Annexe 7) permettant d'estimer le taux d'utilisation de cette échelle. Les questionnaires ont été remplis de façon anonyme par des kinésithérapeutes, neurologues, radiologues interventionnels et réanimateurs de différents hôpitaux (hôpital Laveran à Marseille, hôpital Sainte Anne à Toulon, hôpital de Percy et Bégin à Paris) ainsi que par des membres d'un groupe de recherche sur un réseau social (Facebook).

Nous avons eu 90 réponses exploitables (32 kinésithérapeutes, 21 neurologues, 18 radiologues et 19 réanimateurs).

Voici les principaux résultats :

Connaissez-vous l'échelle NIHSS ?	Oui chez 56 % des kinésithérapeutes	Oui chez 100 % des neurologues	Oui chez 100 % des radiologues	Oui chez 100% des Réanimateurs
A quelle fréquence appliquez-vous cette échelle ?	Souvent chez 5% des kinésithérapeutes	Souvent chez 92% des neurologues	Souvent chez 12 % des radiologues	Souvent chez 85 % des réanimateurs
Utilisez-vous les résultats de cette échelle pour mettre en place votre traitement ?	Oui chez 22% des kinésithérapeutes	Oui chez 100% des neurologues	Oui chez 90% des radiologues	Oui chez 100% des Réanimateurs
Utilisez-vous cette échelle pour prédire le devenir fonctionnel des patients ?	Oui chez 14 % des kinésithérapeutes	Oui chez 90% des neurologues	Oui chez 85% des radiologues	Oui chez 85% des réanimateurs
Utilisez-vous ce score uniquement pour classer la sévérité de l'AVC ?	Oui chez 40 % des kinésithérapeutes	Oui chez 10% des neurologues	Oui chez 15% des radiologues	Oui chez 15% des réanimateurs
A quel moment de la prise en charge réalisez-vous ce score ?	A 7 jours chez 38% des kinésithérapeutes	A l'admission chez 90% des neurologues	A 2 heures chez 80% des radiologues	A l'admission chez 70% des réanimateurs
Selon vous quel serait le devenir fonctionnel à 3 mois d'un homme de 45 ans ayant fait un AVC ischémique avec un score de 5 au NIHSS réalisé à 2 heures de l'AVC et avant toute thérapeutique ?	Bonne réponse chez 75% des kinésithérapeutes	Bonne réponse chez 100% des neurologues	Bonne réponse chez 100% des radiologues	Bonne réponse chez 100% des réanimateurs
Même question pour une femme ?	Bonne réponse chez 5% des kinésithérapeutes	Bonne réponse chez 75% des neurologues	Bonne réponse chez 50% des radiologues	Bonne réponse chez 60% des réanimateurs

Fig. 20 : Résultat du questionnaire

Ces résultats ne sont donnés qu'à titre indicatif et mériteraient une véritable étude plus approfondie et avec une meilleure qualité méthodologique pour être confirmés.

Nous pouvons néanmoins dire que l'échelle NIHSS est très largement utilisée parmi les personnes sondées. Elle est réalisée le plus souvent par les neurologues mais les autres professionnels de santé utilisent largement ce score pour leurs traitements respectifs. D'ailleurs le NIHSS est recommandé par l'HAS pour la prise en charge précoce des AVC (recommandation du 4 mai 2009).

A noter toutefois que les kinésithérapeutes connaissent et utilisent très peu cet outil à l'inverse d'une échelle comparable très largement utilisée : le score ASIA pour les blessés médullaires.

Plusieurs hypothèses à ce constat :

- Le NIHSS est le plus souvent réalisé à l'admission du patient aux urgences et n'est pas réévalué ultérieurement. Or le kinésithérapeute n'intervient pas auprès du patient à ce moment.
- Les dossiers médicaux ne sont parfois pas disponibles ou accessibles
- Le score NIHSS n'est pas toujours rapporté sur ces dossiers patients
- Il existe une idée générale que c'est un score médical seulement utile pour la classification de l'AVC et pour choisir le traitement médical

En conclusion, c'est une échelle très largement utilisée ainsi, il est facile pour le kinésithérapeute de connaître le score NIHSS du patient à l'admission pour pouvoir l'utiliser. De plus le temps moyen pour réaliser ce test est de 6 minutes et 30 secondes, le kinésithérapeute peu donc très facilement évaluer son patient à l'aide de cette échelle.

Le score ainsi obtenu peut-être un NIHSS initial si le patient n'a pas bénéficié de ce test sinon le kinésithérapeute peut juger de l'évolution du score NIHSS voir calculer le delta NIHSS (variation du NIHSS dans le temps) comme le suggère Wouters (2018).

Pour aller encore plus loin, comme le démontre l'étude de Lyden (1994) [26], le kinésithérapeute peut se former à l'utilisation de cette échelle grâce à un entraînement vidéo, fiabilisant de ce fait le NIHSS.

➤ Incidence sur le choix de traitement

Lors de la réalisation de notre BDK le NIHSS peut-être un outil supplémentaire afin de choisir le traitement le plus adéquat pour le patient. Pour être plus précis le score nous permettra de savoir quelles techniques nous allons privilégier voir quels bilans spécifiques nous allons réaliser.

En annexe 6 nous proposons un arbre décisionnel réalisé depuis les résultats obtenus. En fonction du patient, du type d'AVC et du score NIHSS, le kinésithérapeute pourra estimer le devenir fonctionnel du patient et ainsi adapter son traitement.

Par exemple :

- Si le score NIHSS est bas et le devenir fonctionnel à 3 mois favorable la rééducation en phase aigüe sera axée sur le renforcement musculaire et la préparation à la mise en charge du patient. De même un AVC mineur voir modéré nous imposera de faire des tests ayant une meilleure sensibilité afin de détecter et de traiter des conséquences moins facilement visible (BEST test, Fugl Meyer, Berg etc.)
- A l'inverse si le devenir fonctionnel est défavorable la rééducation sera axée sur la prévention des troubles du décubitus, sur l'autonomie au lit, l'utilisation du fauteuil etc.
Les études de Diserens (2011) [27] et Schwartz (2009) [28] illustrent parfaitement ceci. Dans la première étude les auteurs nous montrent que la mobilisation et la sortie précoce du lit permettent de réduire les risques de complications chez les patients ayant un score NIH supérieur à 6. La seconde étude montre l'intérêt de la marche assistée par un robot chez les patients ayant un NIHSS entre 6 et 20.

Ce ne sont que 2 exemples d'études associant un traitement kinésithérapique au score NIHSS du patient. Il serait intéressant de réaliser une revue de littérature listant les différents traitements

recommandés en fonction du score NIHSS. Ce guide de bonne pratique pourrait aider le kinésithérapeute à choisir un traitement spécifique pour le patient.

Ce travail a été réalisé pour le membre supérieur et une application a été créée « via therapy ». Dans cette dernière, l'utilisateur doit rentrer les données cliniques du patient et l'application propose les thérapies en kinésithérapie les plus adaptées en fonction de la littérature. Un travail similaire pourrait être envisagé pour la récupération des membres inférieurs voir le recouvrement fonctionnel global.

La connaissance du devenir fonctionnel du patient peut impliquer deux effets complètement différents :

- Effet Pygmalion ou effet Rosenthal du nom du psychologue ayant expérimenté pour la première fois ce phénomène. L'idée de Rosenthal est que si on croit très fort en quelque chose, les chances qu'elle se produise augmentent. C'est une prophétie autoréalisatrice. Il est utilisé en science de l'éducation mais est transplantable à la santé. L'expérience initiale a été faite sur des souris. On a dit à un groupe d'étudiant que la souris reçue avait été sélectionnée selon son intelligence alors que le choix avait été aléatoire. On a constaté que les souris dites intelligentes avaient au bout d'un certain temps de meilleures performances. La même expérience a été faite avec des enfants. Le fait de faire croire aux enseignants que certains avaient de meilleures capacités intellectuelles augmentait le résultat scolaire de ces derniers. Là encore une simple croyance améliore les performances.

Pour nous, kinésithérapeutes, ces résultats doivent nous alerter. La connaissance du devenir fonctionnel du patient nous poussera t'elle, inconsciemment, à plus solliciter les patients ayant un devenir fonctionnel favorable ou à l'inverse nous fera être moins impliqué avec les patients ayant un devenir fonctionnel défavorable ?

- Cathy Stinear nous donne des éléments de réponse. Cet auteur a mis au point un algorithme permettant de prédire le devenir fonctionnel du membre supérieur chez les patients ayant eu un AVC [29]. Dans une deuxième étude elle a donné les résultats de cet algorithme à un groupe de kinésithérapeutes et pas à d'autres [30]. On constate que le simple fait de savoir le devenir fonctionnel du membre supérieur du patient réduisait le temps d'hospitalisation d'une semaine. Ainsi l'efficacité du traitement mise en place était plus importante quand le kinésithérapeute savait le résultat du test. Il y a donc une amélioration globale de l'efficacité du traitement chez les patients (bon et mauvais devenir fonctionnel confondu). Comme l'auteur n'a pas évalué séparément l'efficacité du traitement chez les patients ayant un devenir fonctionnel favorable et défavorable, il est impossible de savoir si les 2 ont eu une amélioration de l'efficacité du traitement identique, si l'amélioration était plus marquée chez une des 2 catégories de population ou s'il y a eu une amélioration marquée de l'efficacité du traitement chez les patients avec un devenir fonctionnel favorable et une petite diminution de cette même efficacité chez les patients avec un devenir fonctionnel défavorable.

Ceci est très intéressant pour nous et il serait judicieux de prévoir une étude sur les conséquences, sur le devenir fonctionnel du patient, de la connaissance du NIHSS par le kinésithérapeute. Dans cette étude il est important d'avoir des résultats séparés pour les patients avec un bon devenir fonctionnel et pour les patients avec un devenir fonctionnel défavorable.

➤ Choix des dispositifs en subaigüe

Ce score est de plus en plus utilisé pour prévoir et préparer le parcours du patient à la sortie de l'hôpital (centre de rééducation, domicile accompagné 24H/24, hospitalisation à domicile, domicile seul... Plusieurs études intéressantes évaluent ceci : Schlegel (2003) [31], Schlegel (2004) [32], Rundek (2000) [33] ...

4.3 Qualité des preuves et limites de la revue

Les études incluses dans la revue de littérature présentent certaines limitations qui nous poussent à pondérer la qualité des résultats analysés.

La validité prédictive du NIHSS est largement confirmée par cette revue de littérature.

Néanmoins les études nous donnent la sensibilité et la spécificité de l'échelle au niveau du seuil le plus significatif. Or ces 2 valeurs dépendent de la prévalence de la maladie. C'est pour cette raison que nous avons calculé la RV+ et RV-. Toutefois ces 2 valeurs ne sont pas données par les auteurs ce qui constitue une limite importante.

Nous n'avons pas les données suffisantes pour calculer les intervalles de confiance de RV+ et RV-. Ainsi nous pouvons affirmer que le NIHSS a une très bonne validité prédictive mais les différentes conclusions que nous avons formulées doivent être pondérées.

En effet nous avons trouvé des valeurs différentes en fonction du sexe, du territoire atteint, de l'âge et de la sévérité de l'atteinte. Mais comme nous n'avons pas d'intervalle de confiance des RV+ et RV- il est impossible de dire si les différences trouvées sont significatives ou simplement dû au hasard. C'est pour cette raison que 2 méta-analyses ont été réalisés : celle de la sensibilité et celle de la spécificité aux Cutoff Optimal. Même si ces 2 outils mathématiques présentent des limites, leur analyse (grâce aux intervalles de confiances) confirment les tendances observées pour les valeurs des RV+ et RV- et nous suggèrent que la taille d'effet retrouvée n'est pas simplement dû au hasard.

Nous avons pris le parti de prendre le score mRS comme indicateur. A posteriori il sera important de confronter ces résultats avec d'autres indicateurs : IRM, échelle Barthel ...

Pour être complet il est important de préciser les différentes limites identifiées de cette échelle :

- Un score donné peut représenter des déficits fonctionnels complètement différents. Par exemple un score de 2 peut être obtenu par une paralysie faciale modérée ou une aphasie sévère. Hors ces 2 déficits n'entraînent pas les mêmes conséquences dans les activités de la vie quotidienne.
- La réponse aux 3 premiers items fait intervenir le langage : donc un infarctus de l'hémisphère gauche étendu entraîne un score supérieur de 5 points environ à un infarctus de l'hémisphère droit de même taille.
- La sensibilité du NIHSS est limitée par le faible nombre d'items ainsi que la faible gradation de chaque item. L'avantage de ceci est que l'échelle est plus rapidement réalisable et plus reproductible.
- Certains items sont peu reproductibles. Ceci a poussé les scientifiques à développer des échelles modifiées ayant d'excellents taux de corrélation avec le NIHSS (cf. Annexe 2 études exclues car étudient la validité du mNIHSS)

- Chez les AVC très sévères certains items ne peuvent pas être évalués. Hors l'hétéroévaluation par un accompagnant ne peut pas être faite ce qui limite cette échelle pour ces patients [34].
- Le NIHSS est sous-évalué pour les AVC de la circulation postérieure.

4.4 Biais potentiels de la revue

Afin d'évaluer les biais potentiels de la revue et d'avoir un regard critique sur le travail réalisé nous allons utiliser l'échelle R-AMSTAR [35] (Revised A Measurement Tool to Assess Systematic Reviews, ou « un outil de mesure servant à évaluer les revues systématiques »).

La traduction en français est donnée en annexe 8 par l'institut National d'Excellence en Santé et en Services Sociaux (INESS) Québec.

1. Un plan de recherche établi a priori est-il fourni ?

Le protocole de recherche a été réalisé avant la phase de recherche d'études et cette dernière s'est déroulé conformément aux désidératas du protocole. Dans la partie méthode les critères d'inclusion sont décrits et la question de recherche est bien ciblée (question PICO). Toutefois le protocole de recherche n'a pas été publié préalablement (sur PROSPERO par exemple). Score = 3/4.

2. La sélection des études et l'extraction des données ont-ils été confiés à au moins deux personnes ?

La sélection des études et l'extraction des données a bien été réalisé par 2 personnes de façon indépendante. Les désaccords ont été débattu et en cas de doute une 3eme personne a réglé le différend. Score = 4/4.

3. La recherche documentaire était-elle exhaustive ?

La recherche documentaire était exhaustive car 2 sources électroniques ont été consultées, l'horizon temporel et les bases de données interrogées sont données, les mots clefs et la stratégie de recherche est décrite, une recherche manuelle a été faite et les références des études incluses ont été analysées. Pour être complet une partie de la littérature grise n'a pas été consulté par manque de temps (ouvrage en anglais traitant les suites de l'AVC comme par exemple Stroke de J.P. Mohr). Score = 4/4.

4. La nature de la publication (littérature grise, par exemple) était-elle un critère d'inclusion ?

Tous les rapports ont été inclus. Il n'y a pas eu d'exclusion selon le type de publication ou la langue par exemple. L'ensemble des articles trouvé était traduit en anglais. Score = 4/4.

5. Une liste des études (incluses et exclues) est-elle fournie ?

Il y a un tableau résumant les caractéristiques des études incluses ainsi qu'un tableau décrivant les études exclues ainsi que le motif. Le lecteur peut retracer les études incluses et/ou exclues en suivant la méthodologie décrite. Score = 4/4.

6. Les caractéristiques des études incluses sont-elles indiquées ?

Deux tableaux complets des caractéristiques des sujets inclus dans les études ont été réalisés. 2 tableaux simplifiés ont été incorporé dans la revue de littérature afin d'avoir une meilleure visibilité. Score = 4/4.

7. La qualité scientifique des études incluses a-t-elle été évaluée et consignée ?

L'échelle Quips Tools a été utilisé pour évaluer la qualité des études incluses. Comme le préconise le protocole, les études présentant un fort risque de biais ont été exclues. Les résultats et la discussion tiennent compte du score au Quips Tools. Score = 4/4.

8. La qualité scientifique des études incluses dans la revue a-t-elle été utilisée adéquatement dans la formulation des conclusions ?

Les conclusions tiennent compte de la qualité scientifique des différentes études. Néanmoins la revue de littérature n'est pas orientée vers la production de guides de pratique ou de recommandations. Un travail de plus longue durée est envisagé afin de produire une recommandation de bonnes pratiques en kinésithérapie en fonction du score NIHSS. Score = 2/4.

9. Les méthodes utilisées pour combiner les résultats des études sont-elles appropriées ?

Les résultats ont été combinés pour l'AUC Curve, pour la sensibilité ainsi que pour la spécificité. Néanmoins le test objectif d'homogénéité n'a pas été réalisé par manque d'outils informatiques et/ou de connaissances statistiques. Toutefois une appréciation qualitative de l'homogénéité a été donnée. Score = 1/4.

10. La probabilité d'un biais a-t-elle été évalué ?

La revue de littérature ne prend pas en compte les biais de publication ou de l'effet tiroir. Les outils graphiques ou test statistiques n'ont pas été réalisés. Score = 1/4.

11. Les conflits d'intérêt ont-ils été déclarés ?

Les conflits d'intérêts ont été déclarés par l'auteur. Score 4/4.

Conclusion : score de 35/44.

Ainsi cette revue de littérature présente une bonne méthodologie et respecte les grands principes de ce type d'article. Les conclusions formulées peuvent être utilisées par les kinésithérapeutes dans leur pratique quotidienne. Cependant certaines limites principalement dû au manque d'outils statistiques poussent à pondérer les conclusions de la revue.

Nous avons obtenu des résultats similaires en utilisant la ROBIS Tool [36]. Afin d'éviter les redondances, l'ensemble des résultats est en annexe 9.

De plus la rédaction de la revue de littérature a respecté les 27 recommandations des lignes directives PRISMA [37] (pour la rédaction des revues de littérature et de méta-analyses).

5. Conclusion

Cette revue de littérature nous a permis de confirmer que le NIHSS avait une très bonne validité prédictive. Ainsi cette échelle, en plus d'être un outil permettant une classification des patients selon la gravité de l'AVC lors de l'arrivée aux urgences, permet également de prédire le devenir fonctionnel du patient.

La validité et la fiabilité sont les 2 critères qui nous permettent d'évaluer un outil diagnostique ou pronostique. Ils doivent être recherchés avant toute utilisation de l'outil sous peine que les valeurs obtenues ne soient pas en accord avec la réalité.

Mais ce qui est ressorti tout au long de cette revue c'est que dire qu'un score a une bonne validité prédictive n'est pas suffisant et ne nous permet pas d'utiliser ce test correctement. En effet il est important de tenir compte de la population concernée, de la pathologie et du moment de réalisation du score. Concernant le NIHSS on a pu voir que le moment idéal de réalisation est à 24 heures après le traitement et que le seuil de significativité était différent en fonction de la zone atteinte (circulation postérieure ou antérieure, atteinte supra ou infratentorielle) en fonction de l'âge, du sexe, de la sévérité et du type d'AVC. De plus, on a pu constater que pour une échelle quantitative continue, avoir un seuil est fondamental dans la pratique quotidienne.

Le NIHSS est un outil largement utilisé en neurologie mais il est malheureusement mal connu et peu mis à profit par les kinésithérapeutes. Or nous sommes les principaux acteurs du devenir fonctionnel d'un patient à la suite d'un AVC, nous pouvons et devons donc utiliser cette ressource afin de choisir le traitement le plus adéquat. Ce score rentre donc dans cette démarche d'EBP de notre profession. Il est important que nous nous appuyions sur des outils fiables et valides afin de proposer un traitement personnalisé.

Le NIHSS est considéré comme un outil à destination des médecins. Nous avons démontré tout au long de ce travail que cet outil peut être utilisé par le kinésithérapeute pour améliorer sa pratique. Nous pouvons élargir ce constat à l'ensemble des outils utilisés par d'autres professionnels de santé (médecins, infirmiers, podologues, ergothérapeutes, orthophonistes...) que nous devons utiliser pour être plus pertinent dans notre diagnostic et la mise en place du traitement. De plus l'utilisation de ces tests donnera plus de crédibilité à notre profession et nous permettra une meilleure relation interdisciplinaire en parlant le même langage.

Ces conclusions sont également généralisables à l'ensemble des outils utilisés en kinésithérapie afin d'avoir une meilleure pratique EBP. En effet il est important d'asseoir notre pratique sur des indicateurs et des marqueurs objectifs pour que notre diagnostic kinésithérapique, le choix des traitements et l'efficacité de ces derniers ne soient pas empiriques mais concrets et démontrables.

Ce travail de fin d'étude fut réellement formateur pour moi et constitue la conclusion d'un long travail de réflexion et de maturation professionnelle de 4 ans. En effet, en première année, il a commencé par le choix de la thématique (neurologie). Ma deuxième année a été consacrée à un long travail de recherche préliminaire sur les différents facteurs pronostiques du devenir fonctionnel du patient ayant subi un AVC. Choisir le NIHSS est devenu une évidence en troisième année. Quant à ma quatrième année, elle a été dédiée à la recherche et la rédaction de mon mémoire. Durant ces 4 années j'ai acquis les différents outils indispensables à la réalisation de ce travail. Quand je relis mon travail de fin d'étude du diplôme d'état de manipulateur en radiologie, je me rends compte du chemin parcouru et me sens

aujourd'hui prêt à pouvoir utiliser la littérature dans ma pratique quotidienne et à publier à mon tour afin de faire évoluer notre profession.

Cette revue soulève bon nombre de questions qui mériteront dans l'avenir un travail de recherche plus approfondi. En effet il serait intéressant de réaliser une revue de la littérature concernant un score de prédiction clinique incluant le NIHSS et l'âge par exemple. Le but étant de chercher une prédiction plus juste et plus fine du devenir fonctionnel du patient.

En allant plus loin il serait intéressant de faire un guide des bonnes pratiques répertoriant les différentes techniques utilisées par les kinésithérapeutes pour les patients ayant eu un AVC et voir l'intérêt de chacune en fonction du score NIHSS du patient.

Pour finir, à l'image de l'étude de Stinear (2017), il serait pertinent et relativement simple de faire une étude pour voir les conséquences sur le devenir fonctionnel du patient de la connaissance du NIHSS par le kinésithérapeute.

Ce travail m'a confirmé dans l'orientation professionnelle que je souhaite prendre : travail en neurologie, continuer la recherche en kinésithérapie afin de publier et de transmettre aux professionnels et étudiants. En travaillant à l'Hôpital d'Instruction des Armées Sainte Anne j'aurais l'occasion de réaliser ceci et j'espère un jour avoir l'opportunité de transmettre mes connaissances et ma passion aux étudiants au sein de l'Institut de Formation en Masso-Kinésithérapie de Marseille.

Bibliographie

- [1] C. Lecoffre, C. De Peretti, O. Griamud, F. Woimant, Y. Béjot, and V. Olié, "Patients hospitalisés pour accident vasculaire cérébral en France en 2014 et évolutions depuis 2008 selon le type d'AVC," *Rev. Epidemiol. Sante Publique*, vol. 65, p. S8, 2017.
- [2] T. Brott *et al.*, "Measurements of acute cerebral infarction: A clinical examination scale," *Stroke*, vol. 20, no. 7, pp. 864–870, 1989.
- [3] B. C. Meyer, T. M. Hemmen, C. M. Jackson, and P. D. Lyden, "Modified National Institutes of Health Stroke Scale for use in stroke clinical trials : Prospective reliability and validity," *Stroke*, vol. 33, no. 5, pp. 1261–1266, 2002.
- [4] L. B. Goldstein and G. P. Samsa, "Reliability of the National Institutes of Health Stroke Scale," *Stroke*, vol. 28, no. 2, pp. 307–310, 1997.
- [5] J. RANKIN, "Cerebral vascular accidents in patients over the age of 60. III.," *Scott. Med. J.*, vol. 2, no. 6, pp. 254–268, 1957.
- [6] J. C. Van Swieten, P. J. Koudstaal, M. C. Visser, H. Schouten, and J. Van Gijn, "Interobserver agreement for the assessment of handicap in stroke patients," *Stroke*, vol. 19, no. 5, pp. 604–607, 1988.
- [7] C. D. A. Wolfe, N. A. Taub, E. J. Woodrow, and P. G. J. Burney, "Assessment of scales of disability and handicap for stroke patients," *Stroke*, vol. 22, no. 10, pp. 1242–1244, 1991.
- [8] Y. Shinohara, K. Minematsu, T. Amano, and Y. Ohashi, "Modified rankin scale with expanded guidance scheme and interview questionnaire : Interrater agreement and reproducibility of assessment," *Cerebrovasc. Dis.*, vol. 21, no. 4, pp. 271–278, 2006.
- [9] S. Kwon, A. G. Hartzema, P. W. Duncan, and S. M. Lai, "Disability Measures in Stroke : Relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale," *Stroke*, vol. 35, no. 4, pp. 918–923, 2004.
- [10] J. A. Hayden, D. A. Van der Windt, J. L. Cartwright, P. Côté, and C. Bombardier, "Research and reporting methods annals of internal medicine assessing bias in studies of prognostic factors," *Ann Intern Med*, vol. 144, pp. 427–437, 2013.
- [11] H. Delacour, N. François, A. Servonnet, A. Gentile, and B. Roche, "Les rapports de vraisemblance : un outil de choix pour l'interprétation des tests biologiques," *Immuno-Analyse Biol. Spec.*, vol. 24, no. 2, pp. 92–99, 2009.
- [12] M. Garrison and M. R. Johnson, "Fms and Injury," vol. 10, no. 1, pp. 21–28, 2015.
- [13] V. Dubuc, P. Choi, M. D. Hill, and S. B. Coutts, "Can 90-Day NIHSS Be Used for Outcome Assessment in TIA and Minor Stroke Studies ?," *Cerebrovasc. Dis.*, vol. 40, no. 1–2, pp. 97–98, 2015.

- [14] C. Finocchi, M. Balestrino, L. Malfatto, G. Mancardi, C. Serrati, and C. Gandolfo, "National Institutes of Health Stroke Scale in patients with primary intracerebral hemorrhage," *Neurol. Sci.*, vol. 39, no. 10, pp. 1751–1755, 2018.
- [15] V. Mantero *et al.*, "NIHSS as predictor of clinical outcome at 6 months in patients with intracerebral hemorrhage," pp. 94–96, 2019.
- [16] S. Sato *et al.*, "Baseline NIH Stroke Scale Score predicting outcome in anterior and posterior circulation strokes.," *Neurology*, vol. 70, no. 24 Pt 2, pp. 2371–2377, 2008.
- [17] M.-C. Sun and T.-B. Lai, "Initial Stroke Severity Is the Major Outcome Predictor for Patients Who Do Not Receive Intravenous Thrombolysis due to Mild or Rapidly Improving Symptoms," *ISRN Neurol.*, vol. 2011, pp. 1–3, 2011.
- [18] A. Wouters, C. Nysten, V. Thijs, and R. Lemmens, "Prediction of outcome in patients with acute ischemic stroke based on initial severity and improvement in the first 24 h," *Front. Neurol.*, vol. 9, no. MAY, pp. 1–6, 2018.
- [19] Z. Wu *et al.*, "Time-dependence of NIHSS in predicting functional outcome of patients with acute ischemic stroke treated with intravenous thrombolysis," *Postgrad. Med. J.*, vol. 95, no. 1122, pp. 181–186, 2019.
- [20] N. Yang, B. Zhang, and C. Gao, "The baseline NIHSS score in female and male patients and short-time outcome : A study in young ischemic stroke," *J. Thromb. Thrombolysis*, vol. 37, no. 4, pp. 565–570, 2014.
- [21] S. Yoshimura *et al.*, "NIHSS cut point for predicting outcome in supra- vs infratentorial acute ischemic stroke," *Neurology*, vol. 91, no. 18, pp. E1695–E1701, 2018.
- [22] Adams HP Jr, Davis PH, Leira EC, Chang KC, Bendixen BH, Clarke WR, Woolson RF, Hansen MD. Baseline NIH Stroke Scale score strongly predicts outcome after stroke : A report of the Trial of Org 10172 in Acute Stroke Treatment (TOAST)
- [23] Libman R.B, Hansen M.D, "Differences between anterior and posterior Ph," vol. 11040, no. March, p. 2009, 2009.
- [24] J. J. van Middendorp *et al.*, "A clinical prediction rule for ambulation outcomes after traumatic spinal cord injury: a longitudinal cohort study.," Elsevier Ltd, Mar. 2011.
- [25] J. M. Gaspoz and T. Perneger, "les scores de prédiction clinique : définition, dérivation,, validation et limites," *Rev Med Suisse*, vol. 3.21703, 2001.
- [26] P. Lyden *et al.*, "Improved reliability of the NIH stroke scale using video training," *Stroke*, vol. 25, no. 11, pp. 2220–2226, 1994.
- [27] K. Diserens *et al.*, "Early mobilization out of bed after ischaemic stroke reduces severe complications but not cerebral blood flow : A randomized controlled pilot trial," *Clin. Rehabil.*, vol. 26, no. 5, pp. 451–459, 2012.

- [28] I. Schwartz *et al.*, “The Effectiveness of Locomotor Therapy Using Robotic-Assisted Gait Training in Subacute Stroke Patients : A Randomized Controlled Trial,” *PM R*, vol. 1, no. 6, pp. 516–523, 2009.
- [29] C. M. Stinear, P. A. Barber, M. Petoe, S. Anwar, and W. D. Byblow, “The PREP algorithm predicts potential for upper limb recovery after stroke,” *Brain*, vol. 135, no. 8, pp. 2527–2535, 2012.
- [30] C. M. Stinear, W. D. Byblow, S. J. Ackerley, P. A. Barber, and M. C. Smith, “Predicting Recovery Potential for Individual Stroke Patients Increases Rehabilitation Efficiency,” *Stroke*, vol. 48, no. 4, pp. 1011–1019, 2017.
- [31] D. Schlegel *et al.*, “Utility of the NIH stroke scale as a predictor of hospital disposition,” *Stroke*, vol. 34, no. 1, pp. 134–137, 2003.
- [32] D. J. Schlegel, D. Tanne, A. M. Demchuk, S. R. Levine, and S. E. Kasner, “Prediction of hospital disposition after thrombolysis for acute ischemic stroke using the National Institutes of Health Stroke Scale,” *Arch. Neurol.*, vol. 61, no. 7, pp. 1061–1064, 2004.
- [33] T. Rundek *et al.*, “Predictors of resource use after acute hospitalization: The Northern Manhattan Stroke Study,” *Neurology*, vol. 55, no. 8, pp. 1180–1187, 2000.
- [34] Muir KW, Weir CJ, Murray GD, Povey C, Lees KR. Comparison of neurological scales and scoring systems for acute stroke prognosis. *Stroke*. 1996 ;27 :1817–1820
- [35] Kung J, Chiappell F, Cajulis O, Avezova R, Kossan G, Chew L, Maida C. From Systematic Reviews to Clinical Recommendations for Evidence-Based Health Care : Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance 2010
- [36] P. Whiting *et al.*, “ROBIS : A new tool to assess risk of bias in systematic reviews was developed,” *J. Clin. Epidemiol.*, vol. 69, pp. 225–234, 20
- [37] M. Gedda, “Traduction française de dix lignes directrices pour l’écriture et la lecture des articles de recherche,” *Kinesithérapie*. 2015
- [38] S. Lefebvre *et al.*, “Rehabilitation stroke unit : Study of efficiency, comparative not randomized test,” *J. Readapt. Medicale*, vol. 31, no. 1, pp. 6–13, 2011.

Liste des figures

Fig. 1 : campagne de l'INPES

Source : <https://www.santepubliquefrance.fr/presse>

Fig. 2 : anatomie vasculaire cerveau

Source : Netter

Fig. 3 : Hématome intra-parenchymateux pariétal droit envahissant le ventricule droit.

Source : <https://info-radiologie.ch/>

Fig. 4 : A-Diffusion ; B-perfusion ; C-T1 ; D-T2

Source : <https://info-radiologie.ch/>

Fig. 5 : A-Diffusion ; B-TOF ; C-T2 ; D-TOF AVC ischémique sylvien droit

Source : <https://info-radiologie.ch/>

Fig. 6 : Score NIHSS

Source : www.sfm.org/

Fig. 7 : Classification des AVC selon la gravité

Fig. 8 : Score mRS

Source : <https://urgences-serveur.fr/>

Fig. 9 : Question PICO

Fig. 10 : Quips Tools

Source : <https://blogs.biomedcentral.com/>

Fig. 11 : Diagramme de flux PRISMA

Fig. 12 : tableau récapitulatif simplifié de l'extraction des données des études

Fig. 13 : Schéma récapitulatif des biais

Source : Fait depuis le logiciel RevMan

Fig. 14 : Biais des études incluses

Fig. 15 : Tableau récapitulatif simplifié des résultats extraits

Fig. 16 : Méta-analyse de la sensibilité des études

Fig. 17 : Méta-analyse de la spécificité des études

Fig. 18 : Méta-analyse de l'AUC des études

Fig. 19 : schéma récapitulatif

Fig. 20 : Résultat du questionnaire

Lexique mathématique

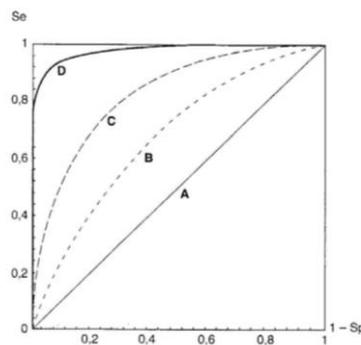
- **Fiabilité** : La fiabilité réfère au degré de constance ou d'exactitude avec lequel un instrument ou technique mesure le concept ou le phénomène qu'il est supposé mesurer (Monkkink et al., 2010). Il existe la fiabilité inter-évaluateur et la fiabilité intra-évaluateur.
Par exemple un pèse-personne est un outil donnant le poids d'un objet ou d'une personne. Pour qu'il ait une bonne fiabilité il doit avoir une constance dans ces mesures. Ainsi le poids donné d'un objet inerte n'ayant pas subi de modification doit être le même lors de 2 mesures consécutives (fiabilité intra-examineur). Tout comme ce poids ne doit pas varier lors de la pesée par 2 examinateurs différents (fiabilité inter-examineur).
- **Validité** : La validité d'un instrument de mesure réfère au degré auquel l'instrument mesure ce qu'il est supposé mesurer, c'est-à-dire le concept évalué. Pour l'exemple précédent, la validité représente la capacité du pèse-personne à donner le bon poids de l'objet. Il y a 3 types de validité : validité de contenu, validité de critère et validité de construit.
- **Validité de contenu** : La validité de contenu est une estimation de la représentativité d'une mesure pour chaque élément d'une mise en situation.
- **Validité de critère** : utilisé si l'on peut comparer les résultats du test à un « gold standard ». Dans l'exemple précédent si l'on a un poids étalon on pourra comparer les résultats donnés par le test à la valeur du poids étalon. Elle peut être concourante (concordance avec le test de référence) ou prédictive (capacité à prédire)
- **Validité de construit** : en cas d'absence de gold standard. On construit une hypothèse à forte probabilité afin d'accepter ou de refuser la mesure.
- **Sensibilité (Se)** : En médecine la sensibilité d'un test est la capacité de ce dernier à être positif chez les patients malades. Dans le tableau de contingence ci-dessous la sensibilité est donnée par la formule : $VP / VP + FN$

	Malade	Non malade
Test positif	VP	FP
Test négatif	FN	VN

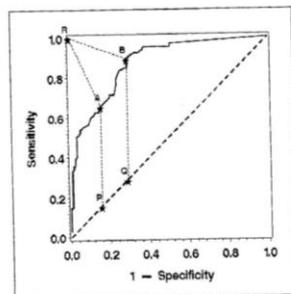
- **Spécificité (Sp)** : La spécificité est la capacité d'un test à être négatif chez les patients non malades. La formule est donc $VN / VN + FP$
- **Validité prédictive positive (VPP)** : représente la probabilité d'être malade en cas de test positif. La formule est : $VP / VP + FP$
- **Validité prédictive négative (VPN)** : représente la probabilité de ne pas être malade en cas de test négatif. La formule est $VN / VN + FN$
- **Rotio (ou rapport) de vraisemblance positif (RV+)** : représente le rapport entre la probabilité d'avoir un test positif chez les sujets malades et la probabilité d'avoir un test positif chez les

sujets sains. La formule est $Se/1-Sp$. Par exemple un $RV+$ de 5 signifie qu'il y a 5 fois plus de chance de présenter un test positif chez les malades que chez les non-malades.

- Rapport de vraisemblance négatif ($RV-$) : représente le rapport entre la probabilité d'avoir un test négatif chez les sujets malades et la probabilité d'avoir un test négatif chez les sujets sains. La formule est : $1-Se/Sp$
- Intérêt $RV+$ et $RV-$: l'intérêt principal repose sur le fait que ces 2 valeurs sont indépendantes de la prévalence de la pathologie dans la population. Ainsi ces 2 rapports représentent mieux les propriétés intrinsèques du test.
- ROC Curve (receiver operating characteristic) : caractéristique de performance (d'un test) ou courbe sensibilité/spécificité. C'est une présentation graphique d'un test permettant de caractériser une variable quantitative continue. Les 2 principaux avantages de cette courbe c'est d'avoir : une représentation complète de la sensibilité et spécificité du test pour toutes ces valeurs et une courbe représentant la propriété intrinsèque du test. Puis cette courbe permet de trouver la valeur du test la plus significative (seuil ou cutoff) ayant la meilleure sensibilité et spécificité. Cette valeur est obtenue grâce au Youden Index ou visuellement grâce au pic de la courbe.



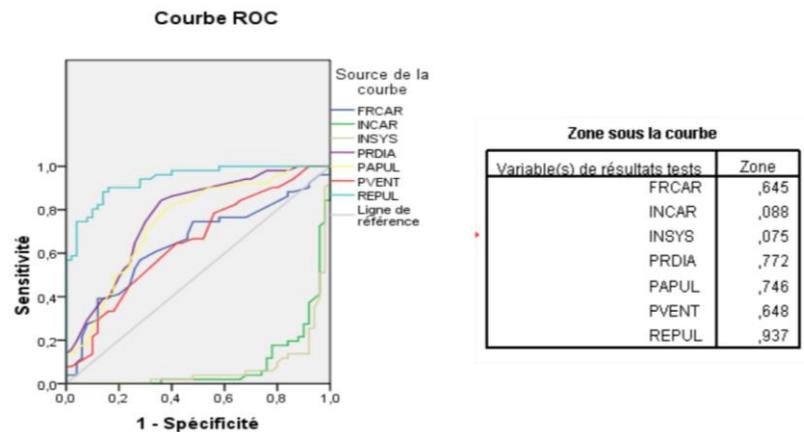
A=Test sans intérêt diagnostique ; B=Mauvais test diagnostique ; C=Test meilleur que B ; D=Bon test diagnostique



Le seuil correspondant

- au point le plus proche de l'idéal (1;1)
- au point le plus loin de la diagonale

- L'AUC ou surface sous la courbe ROC (Area Under Curve) : correspond à l'air sous la courbe ROC Curve. C'est une valeur objective de la qualité intrinsèque du test. Le test parfait a une AUC de 1. Dans l'exemple ci-dessous la courbe bleue REPUL (0,937) représente un test avec d'excellentes propriétés intrinsèques.



- L'indice de Youden : Cet indice mesure la précision d'un outil. Son intérêt est qu'il dépend de la sensibilité, de la spécificité mais pas de la prévalence de la maladie.
 Sa formule est $Y = Se + Sp - 1$.
 Il est compris entre 0 (outil inefficace) et 1 (outil parfait).
 Le calcul de cet indice pour chaque valeur d'une échelle quantitative permet d'avoir la valeur avec le meilleur rapport et donc d'avoir la valeur seuil optimal (cutoff optimal) pour cette échelle. L'indice de Youden est une méthode objective afin de déterminer le seuil d'une échelle.

Source : http://cedric.cnam.fr/~saporta/Sensibilite_specificiteSTA201.pdf

SOMMAIRE DES ANNEXES

Annexe 1	Echelle NIHSS (National Institute of Health Stroke Score)
Annexe 2	Tableaux des études exclues (n=63)
Annexe 3	Tableau complet d'extraction des données
Annexe 4	Quips Tools selon Cochrane.....
Annexe 5	Tableau complet d'extraction des résultats
Annexe 6	Arbre décisionnel.....
Annexe 7	Questionnaire
Annexe 8	R-AMSTAR
Annexe 9	ROBIS : Tool to assess risk of bias in systematic reviews.....

Annexe 1

Echelle NIHSS (National Institute of Health Stroke Score)

1a. Niveau de conscience

Choisir une réponse, même si une évaluation complète est impossible du fait d'une intubation, d'une barrière linguistique, d'un traumatisme oro-trachéal ou d'une autre raison. Le score 3 n'est attribué que si le patient ne réagit par aucun mouvement (en dehors de réflexes de posture) aux stimulations douloureuses.

0	Eveillé, réactivité vive	
1	Non vigilant, mais peut-être éveillé par une simulation mineure (question, ordre)	
2	Non vigilant, nécessite des stimulations répétées pour réagir, ou est inconscient et nécessite des stimulations intenses ou douloureuses pour obtenir des mouvements autres que stéréotypés	
3	On obtient uniquement des réponses réflexes motrices ou végétatives ou le patient est totalement aréactif et flasque	

1b. LOC questions

Demander au patient le mois de l'année en cours, ainsi que son âge. La réponse doit être correcte: les réponses approximatives ne sont pas acceptées. Le score 2 sera attribué aux patients aphasiques ou stuporeux qui ne comprennent pas les questions.

Le score 1 sera attribué aux patients qui ne peuvent répondre du fait d'une intubation, d'un traumatisme oro-trachéal, d'une dysarthrie sévère qu'elle qu'en soit la cause, d'une barrière linguistique, ou de toute autre raison non liée à l'aphasie. Il est important de ne prendre en compte que la première réponse du patient et de ne pas aider le patient de quelque manière que ce soit.

0	Répond aux 2 questions correctement	
1	Ne répond correctement à une seule question	
2	Ne répond correctement à aucune question	

1c. LOC Commandes

Demander au patient d'ouvrir et de fermer les yeux puis de fermer le poing et d'ouvrir la main non parétique. Utiliser 1 autre commande du même type si les mains ne peuvent être utilisées. L'épreuve est considérée comme bonne fait une tentative non équivoque mais ne peut terminer du fait d'une faiblesse. Si le patient ne répond pas à la commande, l'épreuve sera effectuée sur imitation. En cas de traumatisme, d'amputation ou d'un autre empêchement physique, utiliser d'autres ordres simples. Seule la 1ère tentative est prise en compte

0	Exécute correctement les 2 ordres	
1	Exécute un seul ordre sur les deux	
2	N'exécute aucun des 2 ordres	

2. Meilleur regard

Seuls les mouvements horizontaux seront testés. Les mouvements volontaires et réflexes (oculo-céphaliques) seront testés sans épreuve calorique. En cas de déviation conjuguée des yeux ne pouvant être réduite par une activité volontaire ou réflexe, le score 1 sera attribué. En cas de paralysie isolée d'un nerf oculo-moteur (III, IV ou VI), le score 1 sera attribué. L'oculomotricité peut et doit être testée chez tous les patients aphasiques. En cas de traumatisme oculaire, de pansement, de cécité préexistante ou d'une anomalie de l'acuité visuelle ou du champ visuel, l'examineur doit utiliser les mouvements réflexes et attribuer un score. L'existence d'une paralysie partielle de l'oculomotricité pourra être détectée par un déplacement de l'examineur de part et d'autre du patient, en établissant un contact visuel.

0	Normal	
1	Paralysie partielle ce score est attribué quand le regard est anormal pour 1 ou 2 yeux mais sans déviation forcée ni paralysie complète du regard	
2	Déviation forcée ou paralysie complète du regard non vaincue par les mouvements oculo-céphaliques	

3. Champ visuel

Le champ visuel sera testé par confrontation (quadrants supérieurs et quadrants inférieurs), en utilisant le comptage des doigts ou, si nécessaire le clignement à la menace. Les patients doivent être encouragés. Le champ visuel pourra être considéré comme normal si le patient regarde du côté du doigt en mouvement. En cas de cécité unilatérale ou d'énucléation, le champ visuel sera testé sur l'œil indemne.

Le score 1 ne sera attribué que s'il existe une asymétrie nette ou une quadransopie.

Le score 3 sera attribué en cas de cécité complète quelle qu'en soit la raison. Une stimulation bilatérale et simultanée sera effectuée : le score 1 sera attribué en cas d'extinction visuelle et les résultats de cette épreuve seront utilisés pour coter l'item 11.

0	Aucun trouble du champ visuel	
1	Hémianopsie partielle	
2	Hémianopsie complète	
3	Hémianopsie bilatérale	

4 Paralysie faciale

Sur commande ou sur imitation, obtenir du patient de montrer les dents, de lever les sourcils et de fermer les yeux ? En cas de faible réactivité ou de troubles de la compréhension, utiliser une stimulation douloureuse (manœuvre de Pierre Marie et Foix) et tester la symétrie de la grimace. Dans la mesure du possible, ôter les pansements, etc, pouvant gêner l'examen.

0	Mobilité normale symétrique	
1	Paralysie mineure (effacement du sillon nasogénien, asymétrie du sourire)	
2	Paralysie partielle (paralysie de la partie inférieure de la face, totale ou subtotale).	
3	Paralysie complète (faciale supérieure et inférieure).	

5a Motricité du membre supérieur gauche

Le membre examiné est placé dans la position appropriée : membre supérieur en extension (paumes vers le bas) à 90° (en position assise) ou 45° (en position allongée) ; membre inférieur en extension à 30° (toujours examiné en position allongée). Une chute du membre supérieur est prise en compte si elle intervient en moins de 10 secondes au membre supérieur. Le patient aphasique peut être encouragé par stimulation vocale forte ou imitation, mais non par un stimulus nociceptif. Chaque membre est testé successivement, en commençant par le bras non paralysé.

Seulement en cas d'amputation ou de blocage articulaire de l'épaule ou de la hanche, le score 9 sera attribué et l'examineur devra fournir l'explication de ce score.

0	Pas de chute maintien du bras à 90° (ou 45°) pendant 10 secondes.	
1	Chute ; la position ne peut être maintenue pendant 10 secondes, mais le bras ne retombe pas sur le lit ou un autre support.	
2	Le bras ne peut pas atteindre ou maintenir à 90° (ou 45°) tombe sur le lit, mais le patient peut faire un certain mouvement contre la pesanteur.	
3	Absence de mouvement contre la pesanteur ; le bras tombe instantanément.	
4	Absence de mouvement.	
9	Amputation, blocage articulaire, expliquer :	

5b Motricité du membre supérieur droit

0	Pas de chute maintien du bras à 90° (ou 45°) pendant 10 secondes.	
1	Chute ; la position ne peut être maintenue pendant 10 secondes, mais le bras ne retombe pas sur le lit ou un autre support.	
2	Le bras ne peut pas atteindre ou maintenir à 90° (ou 45°) tombe sur le lit, mais le patient peut faire un certain mouvement contre la pesanteur.	
3	Absence de mouvement contre la pesanteur ; le membre inférieur tombe instantanément.	
4	Absence de mouvement.	
9	Amputation, blocage articulaire, expliquer :	

6a Motricité du membre inférieur gauche

Le membre examiné est placé dans la position appropriée : membre supérieur en extension (paumes vers le bas) à 90° (en position assise) ou 45° (en position allongée) ; membre inférieur en extension à 30° (toujours examiné en position allongée). Une chute du membre est prise en compte si elle intervient en moins de 5 secondes au membre inférieur. Le patient aphasique peut être encouragé par stimulation vocale forte ou imitation, mais non par un stimulus nociceptif. Chaque membre est testé successivement, en commençant par le membre non paralysé.

Seulement en cas d'amputation ou de blocage articulaire de l'épaule ou de la hanche, le score 9 sera attribué et l'examineur devra fournir l'explication de ce score.

0	Absence de chute ; la jambe garde la position à 30° pendant 5 secondes.	
1	Chute ; la position ne peut être maintenue pendant 5 secondes, mais la jambe ne retombe pas sur le lit.	
2	La jambe retombe sur le lit dans les 5 secondes, mais le patient peut faire un certain mouvement contre la pesanteur.	
3	Absence de mouvement contre la pesanteur ; le membre inférieur tombe instantanément.	
4	Absence de mouvement.	
9	Amputation, blocage articulaire, expliquer :	

6 b Motricité du membre inférieur droit

0	Absence de chute ; la jambe garde la position à 30° pendant 5 secondes.	
1	Chute ; la position ne peut être maintenue pendant 5 secondes, mais la jambe ne retombe pas sur le lit.	
2	La jambe retombe sur le lit dans les 5 secondes, mais le patient peut faire un certain mouvement contre la pesanteur.	
3	Absence de mouvement contre la pesanteur ; le membre inférieur tombe instantanément.	
4	Absence de mouvement.	
9	Amputation, blocage articulaire, expliquer :	

7 Ataxie des membres

Le but de cet item est de rechercher une atteinte cérébelleuse unilatérale. L'examen est réalisé les yeux ouverts. En cas de déficit visuel, s'assurer que l'épreuve est effectuée dans le champ visuel intact. Les épreuves doigt-nez et talon-genou seront réalisées des deux côtés : une ataxie ne sera prise en compte que si elle ne peut pas être expliquée par un déficit moteur. L'ataxie sera considérée comme absente chez un patient qui ne peut comprendre ou est paralysé. Seulement en cas d'amputation ou de blocage articulaire, le score 9 sera attribué et l'examineur devra fournir l'explication de ce score. En cas de cécité, le test sera effectué en demandant au patient de toucher le nez à partir de la position bras en extension.

0	Absente.	
1	Présente pour un membre.	
2	Présente pour deux membres.	
9	Amputation, blocage articulaire	

8 Sensibilité

Étudier la sensation ou la grimace à la piqûre ou le retrait après stimulation nociceptive chez le patient confus ou aphasique. Seuls les troubles sensitifs liés à l'accident vasculaire cérébral sont pris en compte. L'examen doit porter sur les différentes parties du corps [bras (et non les mains), jambes, tronc et face] pour rechercher un déficit sensitif hémicorporel. Le score 2, "sévère ou total", ne doit être attribué que lorsqu'un déficit sensitif sévère ou total peut être clairement identifié. En conséquence, le score 1 ou 0 sera probablement attribué aux patients stuporeux ou aphasiques. Le score 2 sera attribué au patient atteint d'un accident vertébro-basilaire ayant un déficit sensitif bilatéral. Le score 2 sera attribué, en l'absence de réponse, aux patients quadriplégiques. Le score 2 est attribué aux patients dans le coma (item 1 à 3).

0	Normale	
1	Déficit sensitif discret à modéré ; du côté atteint, la piqûre est moins bien perçue ou non perçue mais le patient est conscient d'être touché.	
2	Déficit sévère à total ; le patient n'a pas conscience d'être touché au niveau de la face, du bras et de la jambe.	

9 Meilleur langage

De nombreuses informations seront obtenues sur le niveau de compréhension lors de la passation des épreuves précédentes. On demandera au patient de décrire la scène se déroulant sur la figure qui lui sera présentée, de nommer les items représentés sur une autre figure et de lire une liste de phrases (voir plus loin les figures). Le niveau de compréhension est évalué à partir de ces épreuves spécifiques, ainsi qu'à partir de la réponse aux consignes données lors des phases précédentes de l'examen. En cas de troubles visuels, on demandera au patient d'identifier des objets placés dans sa main et on évaluera sa production verbale et la répétition. Chez le patient intubé, l'écriture sera évaluée. Le score 3 sera attribué au patient dans le coma (item 1 à 3). L'examinateur doit attribuer un score chez le patient stuporeux ou dont la coopération est limitée ; le score 3 ne doit être attribué que si le patient est mutique et n'exécute aucun ordre simple.

0	Normal, pas d'aphasie.	
1	Aphasie discrète à modérée. Il existe une perte indiscutable de la fluence ou des capacités de compréhension, mais sans limitation significative des idées exprimées, ni de la forme de l'expression. Toutefois, la réduction du discours et/ou de la compréhension, rend la conversation à partir des documents fournis difficile, voire impossible. L'examinateur peut identifier l'image à partir de la réponse du patient.	
2	Aphasie sévère. L'expression est fragmentaire. L'examinateur doit faire des efforts et interroger pour déduire et deviner. L'étendue des informations échangées est limitée. L'examinateur supporte tout le poids de la conversation. Il ne peut identifier les objets à partir des réponses du patient.	
3	Mutisme ; aphasie globale ; pas de langage utile, ni de compréhension du langage oral.	

10 Dysarthrie

Même si l'articulation semble normale, l'expression verbale doit être testée en demandant au patient de lire ou répéter les mots de la liste. L'articulation du langage lors du discours spontané sera testée même chez le patient ayant une aphasie sévère. Le score 9 ne sera attribué qu'en cas d'intubation ou d'autre empêchement " mécanique " ; l'examinateur doit alors fournir la raison de cet empêchement. Le patient ne doit pas être averti que l'examinateur est en train de tester son articulation.

0	Normal	
1	Discrète à modérée ; le patient n'articule pas bien au moins quelques mots et, au pire, peut être compris avec quelques difficultés.	
2	Sévère : le discours du patient est tellement mal articulé qu'il devient inintelligible en l'absence ou hors de proportion avec toute éventuelle aphasie, ou est mutique, anarthrique.	
9	Intubation, autre obstacle mécanique, expliquer :	

11 Extinction ou négligence

Des informations suffisantes peuvent être obtenues lors des épreuves précédentes. Le score 0 est attribué si le patient a un déficit visuel sévère empêchant la recherche d'une extinction visuelle, à condition que l'épreuve à la recherche d'une extinction sensitive soit normale. Le score est 0, si le patient est aphasique, mais donne l'impression de percevoir le stimulus des deux côtés. La présence d'une négligence visuo-spatiale ou d'une anosognosie doit aussi être prise en compte. Dans la mesure où cette anomalie n'est scorée que si elle est présente, cet item est toujours testable.

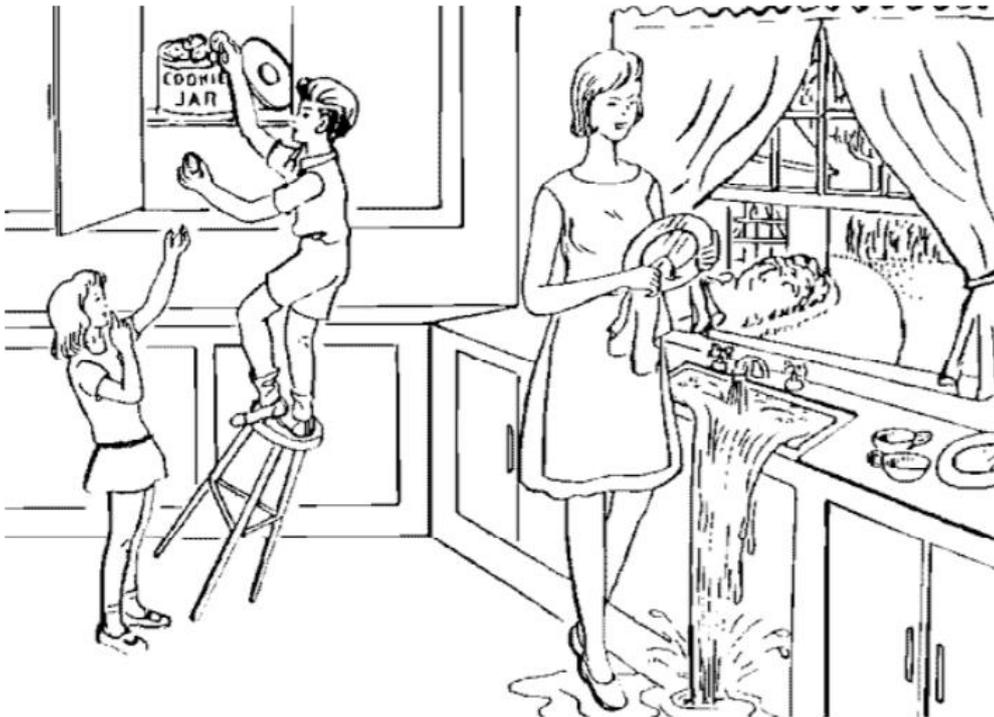
0	Normal	
1	Extinction ou négligence visuelle, tactile, auditive, spatiale ou personnelle à la stimulation bilatérale simultanée dans une des modalités.	
2	Héminégligence sévère ou portant sur plus d'une modalité sensorielle. Ne reconnaît pas sa propre main ou ne s'oriente que vers un côté de l'espace.	

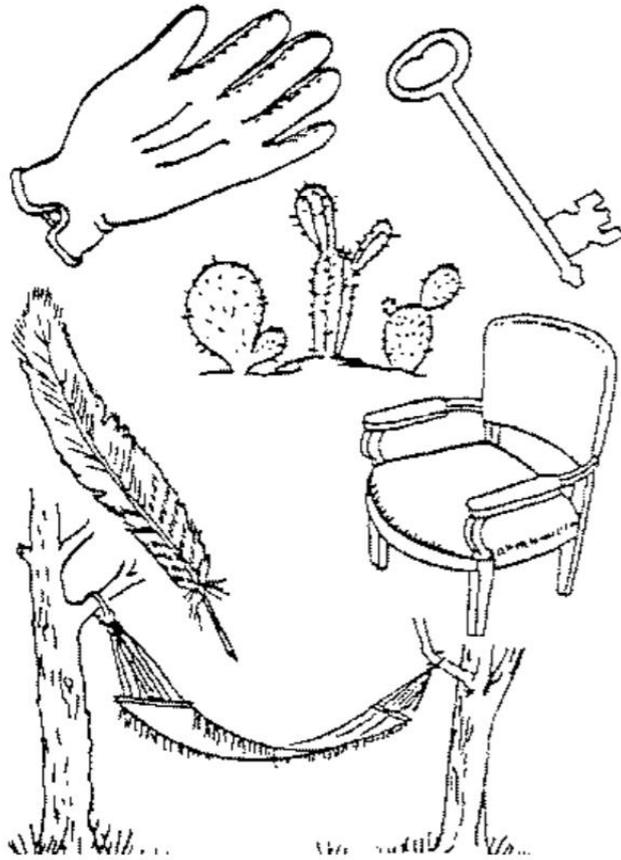
12 Motricité distale

L'examineur soutient l'avant-bras du patient et lui demande d'effectuer une extension complète des doigts. Si le patient ne peut étendre les doigts, l'examineur les place en extension et surveille l'apparition d'une flexion des doigts. Seule la première tentative est prise en compte. Seuls les mouvements lors de cette épreuve sont pris en compte.

0	Normal, pas de flexion après 5 secondes.	
1	Maintien d'une extension après 5 secondes, mais celle-ci est incomplète.	
2	Aucune extension volontaire après 5 secondes.	
9	Amputation, blocage articulaire, expliquer :	

Total





VOUS SAVEZ BIEN

IL TOMBE PAR TERRE

JE VAIS AU TRAVAIL

LE VASE EST DANS LA SALLE À MANGER

JE L'AI ENTENDU HIER SOIR À LA RADIO

MAMAN

TIC TAC

MOITIÉ-MOITIÉ

CINQ

ÉCLABOUSSER

BÉBÉ PLEUREUR

CATÉGORIQUE

Source : <http://www.protocoles-urgences.fr/>

Annexe 2

Tableaux des études exclues (n=63)

Etudes	Raisons d'exclusion
Kwah 2013 ; König 2008 ; Chen 2012 ; Saposnik 2013 ; Weimar 2004	Evaluation d'un indicateur complexe composé du NIHSS et de l'âge du patient (n=5)
Meyer 2009 ; Meyer 2002 ; Kasner 2003 ; Tirschwell 2002 ; Todo 2016 ; Gottesman 2009 ; Olivato 2016	Evaluation validité du NIHSS modifiée (n=7)
Kasner 2006 ; Goldstrein 1997 ; Josephson 2006	Evaluation de la fiabilité du NIHSS (n=3)
Labeit 2018 ; Jeyaseelan 2015	Evaluation de la validité prédictive pour la dysphagie (n=2)
Pego-Perez 2019 ; Askim 2016 ; Young 2005 ; Gray 2009 ; Meyer 2015 ; Celik 2006	Evaluation de la validité du NIHSS en comparaison avec d'autres échelles (n=6)
Appelros 2004 ; Choi 2017 ; Cincura 2009 ; Dawodu 2011 ; Degraha 1999 ; Ghandehari 2013 ; Leira 2013 ; Lindsell 2005 ; Olavarria 2011 ; Peters 2015 ; Siegler 2013 ; Williams 2000	Données non utilisables (n=12)
Goldie 2014 ; Kwakkel 2010	Evaluation de la validité de critère en comparaison avec l'échelle Barthel (n=2)
Baird 2001 ; Fink 2002 ; You 2019 ; Puig 2019 ; P.W. Schaefer 2015	Evaluation de la validité de critère en comparaison avec l'IRM (n=5)
Johnston 2006	Evaluation de la validité de critère en comparaison avec l'échelle GOS (n=1)
Algin 2019 ; Ayis 2013 ; Barber 2004 ; Mark 2004 ; Beckers 2012 ; Brott 1989 ; Cheung 2008 ; Dutta 2017 ; Fonarow 2012 ; Jain 2016 ; Kerr 2012 ; Muir 1996 ; Rundek 2000 ; Satopaa 2017 ; Schlegel 2004 ; Schlegel 2003 ; Strowd 2015 ; Yaghi 2016 ; Zhao 2018 ; Zhong 2016	Evaluation de la validité de critère grâce à d'autres critères : mortalité, décharge, temps d'hospitalisation, risque de réadmission, récurrence (n=20)

Annexe 3

Tableau complet d'extraction des données

Etudes /population	Effectif	Age Moyen	Sexe (H / F)	NIHSS initial médian	Délais réalisation NIHSS	Délais réévaluation /prédiction	Schéma d'étude	Type AVC étudié	Critère de devenir fonctionnel favorable
Dubuc 2015 ¹³	510	N.D.	N.D.	NIHSS ≤3	N.D.	3 mois	Etude cohorte prospective	AVC Ischémique	mRS ≤ 1
Finocchi 2018 ¹⁴	156	75.95 (13,1)	88/68	7 (4-16)	Admission	30 jours 3 mois	Etude cohorte prospective	ICH	mRS ≤ 3
Mantero 2019 ¹⁵	172	75.2 (13)	94/78	10 (4-10)	Admission	6 mois	Etude cohorte prospective	ICH	mRS ≤ 2
Sato 2008 ¹⁶	310	71	202/108	-ant : 8(4-15) -post : 4(2-7)	72h	3 mois	Etude cohorte prospective	- AVC Ischémique antérieur - AVC Ischémique postérieur	mRS ≤ 2
Sun 2011 ¹⁷	49	67.41	30/19	NIHSS≤4 ¹ M= 3	3h	mRS à la décharge	Etude cohorte prospective	AVC Ischémique sans T.I.V	mRS ≤ 2
Wouters 2018 ¹⁸	369 131	70.2 (57,6-78,1) 72.8 (61.5–79)	215/154 71/60	4 (2-7) 9 (7-15)	24h	3 mois	Etude cohorte rétrospective	AVC Ischémique AVC sévère	mRS ≤ 2
Wu 2019 ¹⁹	269	62.68	186/83	6	Admission 2h ;24h ;7jours ; 10jours après traitement	3 mois	Etude cohorte prospective	AVC Ischémique	mRS ≤ 3
Yang 2014 ²⁰	1451	18-50 ans M=42.57	995/456	6	48h	mRS à la décharge	Etude cohorte prospective	AVC Ischémique	mRS ≤ 2
Yoshimura 2018 ²¹	2902	67.80	1829/1073	N.D.	24h et 72h	3 mois	Etude cohorte rétrospective	AVC Ischémique supra et infra tentorial	mRS ≤ 2

Annexe 4

Quips Tools selon Cochrane

Author and year of publication				
Study identifier				
Reviewer				
Biases	Issues to consider for judging overall rating of "Risk of bias"	Study Methods & Comments	Rating of reporting	Rating of "Risk of bias"
Instructions to assess the risk of each potential bias:	These issues will guide your thinking and judgment about the overall risk of bias within each of the 6 domains. Some 'issues' may not be relevant to the specific study or the review research question. These issues are taken together to inform the overall judgment of potential bias for each of the 6 domains.	Provide comments or text excerpts in the white boxes below, as necessary, to facilitate the consensus process that will follow.	Click on each of the blue cells and choose from the drop-down menu to rate the adequacy of reporting as yes, partial, no or unsure.	Click on the green cells; choose from the drop-down menu to rate potential risk of bias for each of the 6 domains as High, Moderate, or Low considering all relevant issues
1. Study Participation	Goal: To judge the risk of selection bias (likelihood that relationship between PF and outcome is different for participants and eligible non-participants).			
Source of target population	The source population or population of interest is adequately described for key characteristics (LIST) .			
Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (number and type used, e.g., referral patterns in health care)			
Recruitment period	Period of recruitment is adequately described			
Place of recruitment	Place of recruitment (setting and geographic location) are adequately described			
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (e.g., including explicit diagnostic criteria or "zero time" description).			
Adequate study participation	There is adequate participation in the study by eligible individuals			
Baseline characteristics	The baseline study sample (i.e., individuals entering the study) is adequately described for key characteristics (LIST) .			
Summary Study participation	The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between PF and outcome.			
2. Study Attrition	Goal: To judge the risk of attrition bias (likelihood that relationship between PF and outcome are different for completing and non-completing participants).			
Proportion of baseline sample available for analysis	Response rate (i.e., proportion of study sample completing the study and providing outcome data) is adequate.			
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.			
Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided.			
Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics (LIST) . There are no important differences between key characteristics (LIST) and outcomes in participants who completed the study and those who did not.			
Study Attrition Summary	Loss to follow-up (from baseline sample to study population analyzed) is not associated with key characteristics (i.e., the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.			

3. Prognostic Factor Measurement	Goal: To judge the risk of measurement bias related to how PF was measured (differential measurement of PF related to the level of outcome).			
<i>Definition of the PF</i>	A clear definition or description of 'PF' is provided (e.g., including dose, level, duration of exposure, and clear specification of the method of measurement).			
<i>Valid and Reliable Measurement of PF</i>	Method of PF measurement is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall). Continuous variables are reported or appropriate cut-points (i.e., not data-dependent) are used.			
<i>Method and Setting of PF Measurement</i>	The method and setting of measurement of PF is the same for all study participants.			
<i>Proportion of data on PF available for analysis</i>	Adequate proportion of the study sample has complete data for PF variable.			
<i>Method used for missing data</i>	Appropriate methods of imputation are used for missing 'PF' data.			
PF Measurement Summary	PF is adequately measured in study participants to sufficiently limit potential bias.			
4. Outcome Measurement	Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).			
<i>Definition of the Outcome</i>	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.			
<i>Valid and Reliable Measurement of Outcome</i>	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).			
<i>Method and Setting of Outcome Measurement</i>	The method and setting of outcome measurement is the same for all study participants.			
Outcome Measurement Summary	Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.			
5. Study Confounding	Goal: To judge the risk of bias due to confounding (i.e. the effect of PF is distorted by another factor that is related to PF and outcome).			
<i>Important Confounders Measured</i>	All important confounders, including treatments (key variables in conceptual model: LIST), are measured.			
<i>Definition of the confounding factor</i>	Clear definitions of the important confounders measured are provided (e.g., including dose, level, and duration of exposures).			
<i>Valid and Reliable Measurement of Confounders</i>	Measurement of all important confounders is adequately valid and reliable (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).			
<i>Method and Setting of Confounding Measurement</i>	The method and setting of confounding measurement are the same for all study participants.			
<i>Method used for missing data</i>	Appropriate methods are used if imputation is used for missing confounder data.			
<i>Appropriate Accounting for Confounding</i>	Important potential confounders are accounted for in the study design (e.g., matching for key variables, stratification, or initial assembly of comparable groups). Important potential confounders are accounted for in the analysis (i.e., appropriate adjustment).			
Study Confounding Summary	Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.			
6. Statistical Analysis and Reporting	Goal: To judge the risk of bias related to the statistical analysis and presentation of results.			
<i>Presentation of analytical strategy</i>	There is sufficient presentation of data to assess the adequacy of the analysis.			
<i>Model development strategy</i>	The strategy for model building (i.e., inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.			
<i>Reporting of results</i>	There is no selective reporting of results.			
Statistical Analysis and Presentation Summary	The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.			

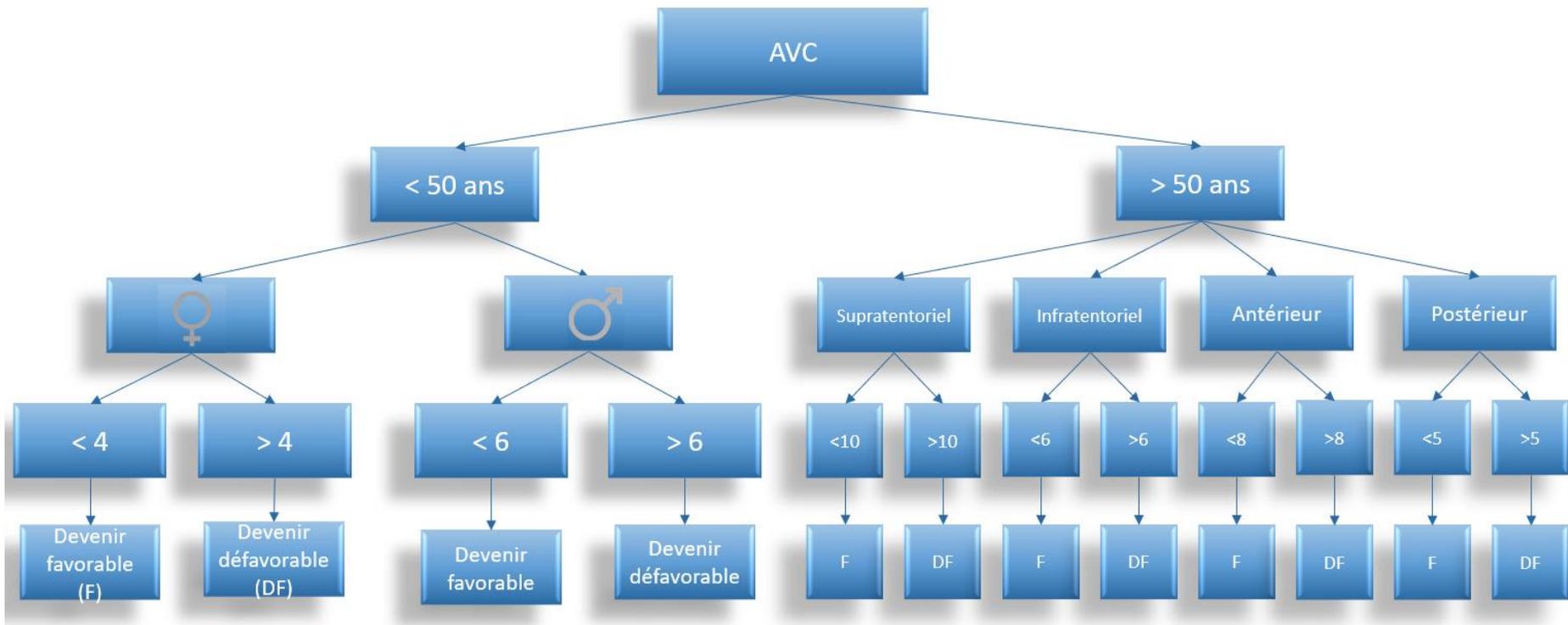
Annexe 5

Tableau complet d'extraction des résultats

Etudes	Sensibilité	Spécificité	RV +	RV-	Valeur prédictive positive	Valeur prédictive négative	AUC ROC Curve	Cut Off NIHSS
Dubuc 2015	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Finocchi 2018	93,5 (à 30 jours) 92,2 (à 3 mois)	82,3 69,6	5,53 3,03	0,07 0,11	81,8 77,9	89,9 81,1	0,914 0,833	7 7
Mantero 2019	87,3 (76,5-94,3)	74,3 (65,1-82,2)	3,40	0,17	66,3 (58,5-73,3)	91 (84-95,1)	0,886	10
Sato 2008	ant : 80 (72-87) post : 84 (74-92)	ant : 82 (72-89) post : 81 (61-93)	ant : 4,44 post : 4,42	ant : 0,24 post : 0,20	ant : 85 (76-91) post : 93 (84-98)	ant : 77 (67-85) post : 64 (45-80)	ant : 0,868 (0,818-0,917) post : 0,867 (0,783-0,951)	ant : 8 Post : 5
Sun 2011	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Wouters 2018	81 (76-85) 71 (58-81)	67 (67-76) 63 (50-75)	2,45 1,92	0,28 0,46	N.D.	N.D.	0,78 (0,72-0,84)	7 10
Wu 2019	1 : 0.51 (0.41-0.61) 2 : 0.73 (0.63-0.81) 3 : 0.83 (0.74-0.90) 4 : 0.85 (0.76-0.91) 5 : 0.85 (0.76-0.91)	1 : 0.84 (0.75-0.91) 2 : 0.75 (0.65-0.84) 3 : 0.65 (0.55-0.74) 4 : 0.76 (0.66-0.84) 5 : 0.80 (0.71-0.87)	1 : 3,19 2 : 2,92 3 : 2,37 4 : 3,54 5 : 4,25	1 : 0,58 2 : 0,36 3 : 0,26 4 : 0,20 5 : 0,19	1 : 36,84 2 : 34,09 3 : 29,82 4 : 38,89 5 : 43,68	1 : 90,57 2 : 93,92 3 : 95,48 4 : 96,65 5 : 96,79	1 : 0,76 2 : 0,79 3 : 0,83 4 : 0,91 5 : 0,92	1 : ≥ 12 2 : ≥ 7 3 : ≥ 5 4 : ≥ 5 5 : ≥ 5
Yang 2014	F : 89 (82-94) H : 79,5 (74,2-84,2)	F : 76 (71-90) H : 86,2 (83,4-88,6)	F : 3,71 H : 5,76	F : 0,14 H : 0,24	F : 95 (92-97) H : 92 (90-94)	F : 57 (49-64) H : 68 (62-73)	F : 0,881 (0,858-0,916) H : 0,906 (0,886-0,923)	F : 4 H : 6
Yoshimura 2018	S : 65 I : 72	S : 73 I : 56	S : 2,41 I : 1,64	S : 0,48 I : 0,50	N.D.	N.D.	S : 0,76 (0,74-0,78) I : 0,69 (0,63-0,75)	S : 10 I : 6

Annexe 6

Arbre décisionnel



Annexe 7

Questionnaire

Profession	Kinésithérapeute	Neurologue	Radiologue	Réanimateur	
Prenez-vous en charge des patients ayant eu un AVC en phase aigüe ?	Oui		Non		
Connaissez-vous l'échelle NIHSS ?	Oui		Non		
A quelle fréquence appliquez-vous cette échelle ?	Jamais	Parfois	Souvent	Systématiquement	
Utilisez-vous les résultats de cette échelle pour mettre en place votre traitement ?	Oui		Non		
Utilisez-vous cette échelle pour prédire le devenir fonctionnel des patients ?	Oui		Non		
Utilisez-vous ce score uniquement pour classer la sévérité de l'AVC ?	Oui		Non		
A quel moment de la prise en charge réalisez-vous ce score ?	A l'admission	A 2 heures	A 24 heures	A 7 jours	A 10 jours
Selon vous quel serait le devenir fonctionnel à 3 mois d'un homme de 45 ans ayant fait un AVC Ischémique avec un score de 5 au NIHSS réalisé à 2 heures de l'AVC et avant toute thérapeutique ?	Favorable		Défavorable		
Même question pour une femme ?	Favorable		Défavorable		

Annexe 8



R – AMSTAR* – ÉVALUATION DE LA QUALITÉ DES REVUES SYSTEMATIQUES

*AMSTAR révisé

AMSTAR : a measurement tool to assess the methodological quality of systematic reviews

Comment fonctionne la liste de contrôle R-AMSTAR?

La liste comporte, dans la colonne gauche, 11 questions sur la qualité de la revue. En fonction des critères présentés dans la colonne droite, attribuer à chaque question un score de 1 à 4. La somme des scores constitue le score de qualité global de la revue systématique.

Éléments AMSTAR

1. Un plan de recherche établi a priori est-il fourni?

La question de recherche et les critères d'inclusion des études doivent être déterminés avant le début de la revue.

Critères

- A. Publication et (ou) inscription du protocole d'étude à l'avance
- B. Description des critères d'inclusion
- C. Question de recherche bien ciblée (critères PICO)

Conditions d'attribution du score

3 critères → 4, 2 → 3, 1 → 2, 0 → 1

Explication A. :

Il doit être dit explicitement que le protocole a été publié ou inscrit, par exemple dans PROSPERO, registre de revues systématiques prospectif et multinational accessible en ligne.
C. La question renferme les critères PICO, soit Population, Intervention (ou exposition), Comparateur (ou témoins) et Résultats (Outcomes).

Score :

Commentaire :

Traduction en français : Institut national d'excellence en santé et en services sociaux www.inesss.qc.ca

Institut national
d'excellence en santé
et en services sociaux
Québec

2. La sélection des études et l'extraction des données ont-ils été confiés à au moins deux personnes?

Au moins deux personnes doivent procéder à l'extraction des données de façon indépendante, et une méthode de consensus doit avoir été mise en place pour le règlement des différends.

<p><i>Critères</i></p> <p>A. Données extraites par au moins deux personnes, de façon indépendante (déclaration explicite ou implicite)</p> <p>B. Énoncé sur le processus de consensus pour le règlement des différends</p> <p>C. Résolution des désaccords entre les personnes ayant extrait les données conformément à la méthode établie (déclaration explicite ou implicite)</p> <p><i>Conditions d'attribution du score</i> 3 critères→4, 2→3, 1→2, 0→1</p>	<p>Score :</p> <p>Commentaire :</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------

3. La recherche documentaire était-elle exhaustive?

Au moins deux sources électroniques doivent avoir été utilisées. Le rapport doit comprendre l'horizon temporel de la recherche et les bases de données interrogées (Central, EMBASE et MEDLINE, par exemple). Les mots clés et (ou) les termes MeSH doivent être indiqués et, si possible, la stratégie de recherche complète doit être exposée. Toutes les recherches doivent être complétées par la consultation des tables des matières de revues scientifiques récentes, de revues de la littérature, de manuels, de registres spécialisés ou d'experts dans le domaine étudié et par l'examen des références fournies dans les études répertoriées.

<p><i>Critères</i></p> <p>A. Au moins deux sources électroniques ont été utilisées.</p> <p>B. L'horizon temporel et les bases de données interrogées sont indiqués.</p> <p>C. Les mots clés et (ou) les termes MeSH sont indiqués et, si possible, la stratégie de recherche est exposée.</p> <p>D. Toutes les recherches sont complétées par la consultation des tables des matières de revues scientifiques récentes, de revues de la littérature, de manuels ainsi que de registres et par l'examen des références fournies dans les études répertoriées.</p> <p>E. Une recherche manuelle a été effectuée dans les revues.</p> <p><i>Conditions d'attribution du score</i> 4 ou 5 critères→4, 3→3, 2→2, 1 ou 0→1</p> <p><i>Explication E. :</i> La recherche manuelle consiste à repérer les revues très pertinentes et à faire une recherche à la main, page par page, de leur contenu afin de relever d'éventuelles études admissibles.</p>	<p>Score :</p> <p>Commentaire :</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------

4. La nature de la publication (littérature grise, par exemple) était-elle un critère d'inclusion?

Les auteurs doivent indiquer s'ils ont recherché tous les rapports, quel que soit le type de publication, ou s'ils ont exclu des rapports (de leur revue systématique) sur la base du type de publication, de la langue, etc.

Critères	Score :
A. Les auteurs indiquent qu'ils ont recherché tous les rapports, quel que soit le type de publication.	Commentaire :
B. Les auteurs indiquent s'ils ont exclu des rapports sur la base du type de publication, de la langue, etc.	
C. « Les articles rédigés dans une langue autre que l'anglais ont été traduits » ou les lecteurs maîtrisaient assez bien la langue du rapport.	
D. Aucune restriction fondée sur la langue ou prise en compte des articles rédigés dans une langue autre que l'anglais	
<i>Conditions d'attribution du score</i> 3 ou 4 critères → 4, 2 → 3, 1 → 2, 0 → 1	

5. Une liste des études (incluses et exclues) est-elle fournie?

Une liste des études incluses et exclues doit être fournie.

Critères	Score :
A. Les études incluses doivent être réunies dans un tableau, une liste ou une figure; une simple liste de références ne suffit pas.	Commentaire :
B. Les études exclues doivent être réunies dans un tableau, une liste ou une figure qui sera intégré à l'article ou à un supplément.	
C. Les raisons de l'exclusion des études sérieusement prises en considération doivent être exposées de manière suffisamment claire.	
D. Le lecteur peut retracer aisément les études incluses et exclues dans la bibliographie, les références ou le supplément de l'article.	
<i>Conditions d'attribution du score</i> 4 critères → 4, 3 → 3, 2 → 2, 1 → 1	
<i>Explication :</i> Les études exclues sont celles qui, après avoir été sérieusement prises en considération sur la foi du titre et (ou) du résumé, ont été rejetées après lecture du corps du texte.	

6. Les caractéristiques des études incluses sont-elles indiquées?

Les données portant sur les sujets qui ont participé aux études originales, les interventions qu'ils ont reçues et les résultats doivent être regroupées, sous forme de tableau, par exemple. L'étendue des données sur les caractéristiques des sujets de toutes les études analysées (âge, race, sexe, données socio-économiques pertinentes, nature, durée et gravité de la maladie, autres maladies, par exemple) doit y figurer.

Critères

- A. Les données portant sur les sujets qui ont participé aux études originales, les interventions qu'ils ont reçues et les résultats sont regroupées, sous forme de tableau, par exemple.
- B. Les auteurs précisent l'étendue des données sur les caractéristiques pertinentes des sujets des études analysées.
- C. L'information fournie semble complète et exacte.

Conditions d'attribution du score

3 critères → 4, 2 → 3, 1 → 2, 0 → 1

Score :

Commentaire :

7. La qualité scientifique des études incluses a-t-elle été évaluée et consignée?

Les méthodes d'évaluation déterminées a priori doivent être indiquées (par exemple, pour les études sur l'efficacité pratique, le choix de n'inclure que les essais cliniques randomisés à double insu avec placebo ou de n'inclure que les études où l'affectation des sujets aux groupes d'étude était dissimulée); pour d'autres types d'études, d'autres critères d'évaluation seront à prendre en considération.

Critères

- A. Les méthodes a priori sont indiquées.
- B. La qualité scientifique des études incluses semble valable.
- C. Le niveau de preuve est exposé, dûment reconnu ou pris en considération.
- D. La qualité des preuves est évaluée ou classée en fonction d'outils d'évaluation de la preuve.

Conditions d'attribution du score

4 critères → 4, 3 → 3, 2 → 2, 1 ou 0 → 1

Explications D. :

Un outil d'évaluation de la preuve est un instrument qui sert à établir le niveau de preuve.
Ex. : l'outil GRADE (*Grading of Recommendations Assessment, Development and Evaluation*).

Score :

Commentaire :

8. La qualité scientifique des études incluses dans la revue a-t-elle été utilisée adéquatement dans la formulation des conclusions?

Les résultats de l'évaluation de la rigueur méthodologique et de la qualité scientifique des études incluses doivent être pris en considération dans l'analyse et les conclusions de la revue, et formulés explicitement dans les recommandations.

<p><i>Critères</i></p> <p>A. Les auteurs ont tenu compte de la qualité scientifique dans l'analyse et les conclusions de la revue.</p> <p>B. La qualité scientifique est formulée explicitement dans les recommandations.</p> <p>C. Les conclusions sont orientées vers la production de guides de pratique.</p> <p>D. L'énoncé de consensus clinique laisse entrevoir la révision ou la confirmation des recommandations de pratique.</p> <p><i>Conditions d'attribution du score</i> 4 critères→4, 3→3, 2→2, 1 ou 0→1</p>	<p>Score :</p> <p>Commentaire :</p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------

9. Les méthodes utilisées pour combiner les résultats des études sont-elles appropriées?

Si l'on veut regrouper les résultats des études, il faut effectuer un test d'homogénéité afin de s'assurer qu'elles sont combinables (chi carré ou I^2 , par exemple). S'il y a hétérogénéité, il faut utiliser un modèle d'effets aléatoires et (ou) vérifier si la nature des données cliniques justifie la combinaison (la combinaison est-elle raisonnable?).

<p><i>Critères</i></p> <p>A. Les auteurs exposent les critères à partir desquels ils ont déterminé que les études analysées étaient assez semblables pour être combinées.</p> <p>B. Dans le cas des résultats regroupés, les auteurs ont fait un test d'homogénéité pour s'assurer que les études étaient combinables.</p> <p>C. Les auteurs ont pris acte du caractère hétérogène (ou non) des études.</p> <p>D. S'il y a hétérogénéité, les auteurs ont utilisé un modèle d'effets aléatoires et (ou) vérifié si la nature des données justifiait la combinaison.</p> <p>E. S'il y a homogénéité, les auteurs exposent la justification ou le test statistique.</p> <p><i>Conditions d'attribution du score</i> 4 ou 5 critères→4, 3→3, 2→2, 1 ou 0→1</p>	<p>Score :</p> <p>Commentaire :</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------

10. La probabilité d'un biais de publication a-t-elle été évaluée?

Une évaluation du biais de publication doit comprendre une association d'outils graphiques (diagramme de dispersion des études ou autre test) et (ou) des tests statistiques (test de régression d'Egger, par exemple).

<i>Critères</i> A. Prise en compte du biais de publication ou de l'effet tiroir B. Outils graphiques (diagramme de dispersion des études, par exemple) C. Tests statistiques (test de régression d'Egger, par exemple) <i>Conditions d'attribution du score</i> 3 critères → 4, 2 → 3, 1 → 2, 0 → 1	Score : Commentaire :
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------

11. Les conflits d'intérêts ont-ils été déclarés?

Les sources possibles de soutien doivent être déclarées, tant pour la revue systématique que pour les études qui y sont incluses.

<i>Critères</i> A. Présentation des sources de soutien B. Absence de conflit d'intérêts – On est ici dans la subjectivité; peut-être faudra-t-il y aller par déduction ou fouiller quelque peu. C. Prise en compte ou énoncé des sources de soutien ou des conflits d'intérêts dans les principales études incluses <i>Conditions d'attribution du score</i> 3 critères → 4, 2 → 3, 1 → 2, 0 → 1	Score : Commentaire :
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------

Score de qualité maximal : 44	Score de qualité :
-------------------------------	--------------------

PEROSH OSH Evidence Methods (Partnership for European Research in Occupational Safety and Health)

©Kung et al. From Systematic Reviews to Clinical Recommendations for Evidence-Based Health Care: Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance. The Open Dentistry Journal, 2010, 4: 84-91.

Annexe 9

ROBIS : Tool to assess risk of bias in systematic reviews

Source : <https://www.bristol.ac.uk/>

Phase 1 : Assessing relevance (Optional)

ROBIS is designed to assess the risk of bias in reviews with questions relating to interventions, aetiology, diagnosis and prognosis. State your overview/guideline question (target question) and the question being addressed in the review being assessed :

Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

For prognostic reviews :

Patients : **Patient ayant fait un AVC**

Outcome to be predicted : **Validité prédictive**

Intended use of model : **Classification et prédiction du devenir fonctionnel du patient**

Intended moment in time : **test réalisé à l'admission, à 24 heures ou à 7 jours**

Does the question addressed by the review match the target question ? **YES**

Phase 2 : Identifying concerns with the review process

DOMAIN 1 : STUDY ELIGIBILITY CRITERIA

Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified :

1.1 Did the review adhere to pre-defined objectives and eligibility criteria ? **Y**

1.2 Were the eligibility criteria appropriate for the review question ? **Y**

1.3 Were eligibility criteria unambiguous ? **Y**

1.4 Were any restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured) ? **Y**

1.5 Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data) ? **PY**

Concerns regarding specification of study eligibility criteria : **LOW**

DOMAIN 2 : IDENTIFICATION AND SELECTION OF STUDIES

Describe methods of study identification and selection (e.g. number of reviewers involved) :

2.1 Did the search include an appropriate range of databases/electronic sources for published and unpublished reports ? **Y**

2.2 Were methods additional to database searching used to identify relevant reports ? **PY**

2.3 Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible ? **PY**

2.4 Were restrictions based on date, publication format, or language appropriate ? **Y**

2.5 Were efforts made to minimise error in selection of studies ? **PY**

Concerns regarding methods used to identify and/or select studies : **LOW**

DOMAIN 3 : DATA COLLECTION AND STUDY APPRAISAL

Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias :

3.1 Were efforts made to minimise error in data collection ? **Y**

3.2 Were sufficient study characteristics available for both review authors and readers to be able to interpret the results ? **PY**

3.3 Were all relevant study results collected for use in the synthesis ? **PY**

3.4 Was risk of bias (or methodological quality) formally assessed using appropriate criteria ? **Y**

3.5 Were efforts made to minimise error in risk of bias assessment ? **Y**

Concerns regarding methods used to collect data and appraise studies rationale for concern : **LOW**

DOMAIN 4 : SYNTHESIS AND FINDINGS

Describe synthesis methods :

4.1 Did the synthesis include all studies that it should ? **Y**

4.2 Were all pre-defined analyses reported or departures explained ? **PY**

4.3 Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies ? **Y**

4.4 Was between-study variation (heterogeneity) minimal or addressed in the synthesis ? **PN**

4.5 Were the findings robust, as demonstrated through funnel plot or sensitivity analyses ? **Y**

4.6 Were biases in primary studies minimal or addressed in the synthesis ? **PY**

Concerns regarding the synthesis and findings : **LOW**

Phase 3 : Judging risk of bias

Summarize the concerns identified during the Phase 2 assessment :

1. Concerns regarding specification of study eligibility criteria : **Tous les critères d'inclusion des études étaient connus et ont été respectés.**

2. Concerns regarding methods used to identify and/or select studies : **2 auteurs ont sélectionné les études. Un troisième a permis de faire un choix en cas d'opinions contraires.**

3. Concerns regarding methods used to collect data and appraise studies : **Les résultats de l'étude Dubuc 2015 n'ont pas pu être intégré car l'auteur ne les a pas fournis.**

4. Concerns regarding the synthesis and findings : **Les principaux risque identifié est l'homogénéité des études incluses dans les méta-analyses et l'absence de résultats quantitatifs sur ce paramètre.**

RISK OF BIAS IN THE REVIEW

Describe whether conclusions were supported by the evidence :

A. Did the interpretation of findings address all of the concerns identified in Domains 1 to 4 ? **Y**

B. Was the relevance of identified studies to the review's research question appropriately considered ? **Y**

C. Did the reviewers avoid emphasizing results on the basis of their statistical significance ? **PY**

Risk of bias in the review RISK : **LOW**

Abstract

Introduction : Le NIHSS est une échelle très largement utilisée. Sa fiabilité et sa validité ont été étudiés dans de nombreuses études. Toutefois aucune revue de littérature n'a été réalisée sur la validité prédictive du NIHSS ainsi que sur les différentes subtilités en fonction de la population et du type d'accidents vasculaires cérébraux (AVC). Le but de cette revue de littérature est donc d'évaluer la validité prédictive du NIHSS.

Méthodes : Pubmed, Science Direct et la littérature grise ont été consultés. Les mots clefs furent : «AVC», «NIHSS», «devenir fonctionnel» et «modified Rankin Scale (mRS)». La méthodologie PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) a été utilisée. La sélection des études fut faite par 2 lecteurs et les désaccords furent tranchés par un troisième lecteur expert. L'extraction des données ainsi que les résultats furent synthétisées dans 2 tableaux. La qualité des études incluses fut évaluée par le Quips tool.

Résultats : Sur les 6240 études obtenues 9 correspondaient aux critères d'inclusion et ont données lieu à une analyse qualitative. Une méta-analyse a été réalisée pour l'aire sous la courbe ROC Curve (AUC) (4 études) et pour la sensibilité et la spécificité (5 études). La validité prédictive du NIHSS est très bonne pour les AVC hémorragiques (AUC comprise entre 0.833 et 0.914) avec une valeur seuil (Cutoff) de 7 pour une prédiction à 1 mois ou 3 mois et de 10 pour une prédiction à 6 mois. Pour les AVC antérieurs et postérieurs la validité prédictive est similaire (AUC de 0.868 et 0.867) alors que le Cutoff est de 8 et 5 respectivement. Pour les AVC sévères l'AUC est plus basse (0.78) et le Cutoff de 10. Pour les AVC chez les patients jeunes, l'AUC est plus élevée (0.881 et 0.906). De plus dans cette population l'AUC est plus élevée chez les hommes (0.906 contre 0.881) et le Cutoff est de 6 pour les hommes et de 4 pour les femmes. La validité prédictive est meilleure pour les AVC supratentoriels (0.76 contre 0.69) et le Cutoff est de 10 pour les AVC supratentoriels et de 6 pour les AVC infratentoriels.

Conclusion : Le NIHSS a une très bonne validité prédictive dans son ensemble. Néanmoins la composition de cette échelle fait que cette validité est variable selon le sexe, la topographie de l'AVC, l'âge et la sévérité initiale de l'atteinte. Pour simplifier l'utilisation et l'interprétation du NIHSS un schéma récapitulatif ainsi qu'un arbre décisionnel ont été proposés.

Mots clefs : AVC ; NIHSS ; mRS ; Validité prédictive ; devenir fonctionnel

Abstract

Background and Purpose : The NIHSS is a scale used often. Its reliability and validity has been studied in numerous studies. However, no literature review has been carried out on the predictive validity of the NIHSS as well as the different characteristics depending on the population and the type of stroke.

The purpose of this systematic review is to assess the predictive validity of the NIHSS.

Methods : Pubmed, Science Direct and grey literature were consulted. The keywords were: "stroke", "NIHSS", "outcome" and "modified Rankin Scale (mRS)". The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology was used. The studies were selected by 2 readers and disagreements were resolved by a third expert reader. Data extraction and results were summarized in 2 tables. The quality of the included studies was assessed by the Quips tool.

Results : Of the 6240 studies obtained 9 met the inclusion criteria and gave rise to a qualitative analysis. A meta-analysis was performed for the area under the ROC Curve (AUC) (4 studies) and for the sensitivity and the specificity (5 studies). The predictive validity of the NIHSS is very good for hemorrhagic strokes (AUC between 0.833 and 0.914) with a Cutoff of 7 for a prediction at 1 month and 10 for a prediction at 3 or 6 months. For anterior and posterior strokes the predictive validity is similar (AUC of 0.868 and 0.867) while the Cutoff is 8 and 5 respectively. For severe stroke, the AUC is lower (0.78) and the Cutoff is 10. For stroke in young patients, the AUC is higher. In addition, in this population the AUC is high for men (0.906 against 0.881) and the Cutoff is 6 for men and 4 for women. Predictive validity is better for supratentorial strokes (0.76 versus 0.69) and the Cutoff is 10 for supratentorial strokes and 6 for infratentorial strokes.

Conclusion : The NIHSS has very good predictive validity as a whole. However, the composition of this scale means that its validity is flexible depending on the gender, the topography of the stroke, the age and the initial severity of the disease. To simplify the use and interpretation of the NIHSS a summary diagram as well as a decision tree was proposed.

Keywords : NIHSS ; mRS ; Validity ; Outcome ; prognosis ; prognostic ; prediction