



Fiabilité et validité de l'échographie lors du diagnostic kinésithérapique dans le cadre de l'évaluation du contrôle moteur chez le patient lombalgique non spécifique : une étude préliminaire



Reliability and validity of a qualitative ultrasound imaging assessment in patients with motor control impairment and non-specific low back pain: A preliminary study

Vincent Ladeveze^a
Sébastien Martin^b
Patrice Piette^c
Fares Zemirline^d
Anthony Demont^b

^a104, Rue Bobillot 75013 Paris, France

^b6, Rue Pétreille 75009 Paris, France

^c4, ruelle du presbytère 35190 Saint-Pern, France

^d29, rue Damesme 75013 Paris, France

Reçu le 16 juin 2018 ; reçu sous la forme révisée le 18 décembre 2018 ; accepté le 18 février 2019

RÉSUMÉ

Contexte. – Nombre d'études montrent l'intérêt de repérer les patients lombalgiques non spécifiques avec un déficit de contrôle moteur afin de les traiter spécifiquement. Mais, peu d'outils fiables, valides et utilisables quotidiennement sont disponibles. Le but de cette étude préliminaire était d'évaluer la fiabilité et la validité d'une évaluation qualitative échographique.

Méthode. – Une étude de fiabilité a été conduite sur quatorze patients lombalgiques (dont 5 atteints d'un déficit de contrôle moteur). Des critères d'évaluation qualitative ont été sélectionnés pour les multifides et le transverse de l'abdomen. La validité a été évaluée en calculant les coefficients kappa entre chaque critère et le résultat de la batterie de tests de Luomajoki.

Résultats. – Sur les dix-neuf critères testés, deux étaient significativement corrélés aux tests de Luomajoki pour les multifides et un pour le transverse ($k > 0,41$ et $p < 0,05$). La fiabilité inter-examineur était moyenne à bonne.

Conclusion. – Cette étude est une première étape pour la validation de critères échographiques qualitatifs. Des études de plus grandes ampleurs sont encore nécessaires.

Niveau de preuve 4.

© 2019 Elsevier Masson SAS. Tous droits réservés

SUMMARY

Background. – Several studies highlight the interest in identifying patients with motor control impairment in non specific low back pain. But, few reliable, valid and usable tools in daily practice are available. The aim of this preliminary study was to evaluate the reliability and validity of a qualitative ultrasound imaging assessment.

Methods. – A reliability study was conducted on fourteen patients (5 with motor control impairment). Qualitative evaluation criteria were selected for the multifidus and the transversus

MOTS CLÉS

Contrôle moteur
Échographie
Fiabilité
Lombalgie commune
Validité

KEYWORDS

Motor control impairment
Qualitative ultrasound
imaging
Reliability
Low back pain
Validity

Auteur correspondant :

V. Ladeveze,
104, Rue Bobillot 75013 Paris,
France
Adresse e-mail :
vincent.ladeveze@gmail.com

abdominis muscle. The validity of ultrasound imaging assessment was established by calculating Cohen's kappa between each criterion and the Luomajoki battery of tests.

Results. – Nineteen criteria were tested. For the multifidus muscle, two criteria were correlated with Luomajoki's tests and one was correlated for the transversus muscle ($k > 0.41$ and $p < 0.05$). The inter-examiner reliability was medium to good.

Conclusion. – This is a first step in the validation of the use of ultrasound imaging. Further studies on a larger scale population are necessary.

Level of evidence 4.

© 2019 Elsevier Masson SAS. All rights reserved

INTRODUCTION

Depuis 2015, les kinésithérapeutes sont habilités à pratiquer l'échographie [1]. Avec une forte prévalence au sein de la population française, la prise en charge de la lombalgie non spécifique (LNS) est une réalité quotidienne du kinésithérapeute [2]. L'échographie est un outil fiable et valide pour évaluer la modification de l'activité musculaire et l'épaississement des muscles [3]. Il serait, par ailleurs, possible d'évaluer le contrôle moteur en qualifiant un schéma d'activation non optimal du muscle transverse de l'abdomen (TrA) [4] ou en identifiant une diminution de volume de la section des muscles multifides (MF) [5]. D'après la littérature, ces muscles seraient les plus impliqués dans la lombalgie [6]. Selon une revue systématique de 2014 [7], le lien entre épaisseur musculaire et lombalgie demeure incertain (pour les MF) voire inexistant (pour le TrA). C'est pourquoi il semble opportun de s'intéresser au comportement global du muscle (schéma d'activation optimal) lors de sa contraction plutôt qu'à la seule mesure de la force musculaire (épaississement). Cependant, la fiabilité et la validité de l'échographie pour une évaluation qualitative de ces muscles restent à établir pour mettre en évidence un déficit de contrôle moteur. Nous proposons donc d'établir une validité de construit [8] afin de valider l'échographie pour l'évaluation qualitative du contrôle moteur chez le patient LNS. Pour être valide l'examen échographique devra refléter une discrimination (un déficit de contrôle moteur) déjà identifiable par les tests de Luomajoki qui feront office de test de référence [9].

L'objectif principal de cette étude préliminaire était d'évaluer la fiabilité inter-examineur et une validité de construit pour des critères échographiques qualitatifs.

La question de recherche peut être formulée ainsi : existe-t-il des critères échographiques valides pour détecter un déficit de contrôle moteur chez les patients LNS ? Les objectifs secondaires étaient d'apporter des informations sur l'intérêt de l'échographie dans le diagnostic kinésithérapique et de proposer un protocole d'évaluation clinique utilisable quotidiennement pour les kinésithérapeutes disposant d'un échographe ou souhaitant s'équiper.

MATÉRIEL ET MÉTHODE

Population

L'étude s'est déroulée dans un cabinet libéral de kinésithérapie du 9^e arrondissement de Paris. Les patients ont été

recrutés parmi la clientèle du cabinet et ont été inclus dans l'étude par ordre d'arrivée. Le recrutement était basé sur la prescription médicale ("lombalgie", "douleur lombaire", etc.) et sur les critères définis ci-après. Afin que les deux évaluateurs puissent réaliser leurs examens en aveugle, seuls les patients débutants une rééducation étaient recrutés. Le kinésithérapeute chargé du recrutement ne faisait pas partie des deux examinateurs et recrutait les patients lombalgiques au sein de sa propre clientèle.

Les critères d'inclusion peuvent être listés de la manière suivante : la présence d'une lombalgie commune diagnostiquée par un médecin, un âge entre 18 et 65 ans et un consentement écrit du patient.

Les critères de non inclusion, eux comprennent : la présence de drapeaux rouges [10] ou de pathologies graves du rachis (infectieuses, tumorales, etc.), la présence d'un déficit moteur des membres inférieurs, un antécédent de chirurgie rachidienne, une douleur en dessous du genou et une grossesse déclarée.

Le recueil des données avait été planifié avant que le test de référence (tests de Luomajoki) et le test étudié (échographie) ne soient effectués. Il s'agit d'une étude prospective.

Méthodes des tests

Déroulement des tests

Les deux évaluations étaient réalisées successivement le même jour. À son arrivée au cabinet le patient devait lire la feuille d'informations, signer le formulaire de consentement et remplir le questionnaire d'incapacité Oswestry [11]. Le but, la durée et le déroulement de l'étude avaient été expliqués lors de la prise de rendez-vous téléphonique. Le patient avait donné son accord préalable au téléphone. La durée du bilan kinésithérapique comprenant les tests de CM de Luomajoki, était de l'ordre de 30 mn. L'examen échographique durait environ 20 mn. Le patient devait prévoir 1 h en tout le jour du rendez-vous.

L'ordre de passage, du test de référence (tests de CM) et du test étudié (échographie), était randomisé par blocs de deux pour éviter un biais potentiel lié à une éventuelle fatigue ou un échauffement musculaire lors du second test. La randomisation de l'ordre de passage avait été prévue à l'avance et réalisée sous Excel via la fonction RANDOM.

Le patient enchaînait alors les deux évaluations qui étaient réalisées par deux examinateurs différents. Les résultats des tests de CM et de l'échographie n'étaient communiqués qu'après la fin des deux examens.

Pour s'assurer de la reproductibilité et du respect du protocole, chaque examinateur disposait d'une version papier

du protocole, d'une check-list bilan ou échographie et d'une feuille de résultat bilan ou échographie.

Test de référence (tests de contrôle moteur)

Lors de l'évaluation clinique, le kinésithérapeute effectuait un bilan de lombalgie standardisé (bilan lombaire IAOM-US®). Après ce bilan, le patient réalisait les tests de Luomajoki [9,12,13]. Il y avait six tests à réaliser l'un après l'autre. À chaque test, 1 point était attribué si le test est positif, 0 si le test était négatif. À la fin des six tests, l'examineur notait le score total sur 6. Si le score était supérieur ou égal à 2, le patient était considéré comme ayant un déficit de contrôle moteur. Parmi les tests, deux évaluent le contrôle moteur en flexion, deux en extension, un teste le contrôle moteur alternativement en flexion et en extension, un dernier évalue l'inclinaison et le contrôle de la rotation [9]. Chaque test dure environ 2 mn et l'examen total se réalise en 10 mn environ. L'évaluateur qui réalisait les bilans était un kinésithérapeute expérimenté (six années de pratique).

Test étudié (échographie)

De nombreuses publications mettent en avant le rôle prépondérant des MF et du TrA dans la lombalgie [6,14–20]. Afin de construire une évaluation applicable en pratique clinique quotidienne, nous avons choisi de cibler ces deux muscles.

L'examen échographique consistait à prendre des images à des moments clés, en mode B (Brillance) afin de pouvoir observer l'épaississement et en mode TM (Temps/Mouvement) pour observer le schéma d'activation (Fig. 1). Ces deux modes sont couramment utilisés en échographie musculo-squelettique [3,21] ; le mode B donne une visualisation des structures en 2D alors que le mode TM permet une visualisation du mouvement de la structure en fonction du temps. Aucune mesure centimétrique n'a été réalisée. L'échographe était positionné de façon à ce que le patient ne puisse pas voir l'image affichée [3]. Pour chaque mesure, l'examineur réalisait trois clichés. Les positions et les manœuvres demandées au patient étaient standardisées (Fig. 2) : l'*abdominal drawing-in maneuver* (ADIM) pour le TrA [22-24] et l'élévation du membre inférieur controlatéral en procubitus pour les MF [5]. Pour le TrA, l'examineur effectuait les mesures en fin d'inspiration et en fin d'expiration. La sonde d'échographie était placée transversalement sur une ligne verticale entre la douzième côte et l'épine iliaque antéro-supérieure (EIAS) en regard du nombril [3]. Pour les MF, le praticien effectuait une mesure au repos puis une en contraction en demandant au patient de décoller la jambe controlatérale de la table d'examen. La sonde d'échographie était placée transversalement en regard de L5 [5].

L'échographe utilisé était un modèle Edge® de la marque Sonosite®. La sonde était linéaire à haute fréquence (13–17 MHz). Le mode choisi était "musculosquelettique", la profondeur 40 mm et le gain ajusté manuellement par l'examineur. Au total, l'évaluation échographique comportait trente-six clichés (18 par muscle, 9 de chaque côté). L'examineur réalisait les clichés puis exportait les images sur une clé USB. Il interprétait ensuite les images en renseignant le tableau de résultats. Pour chaque critère échographique,

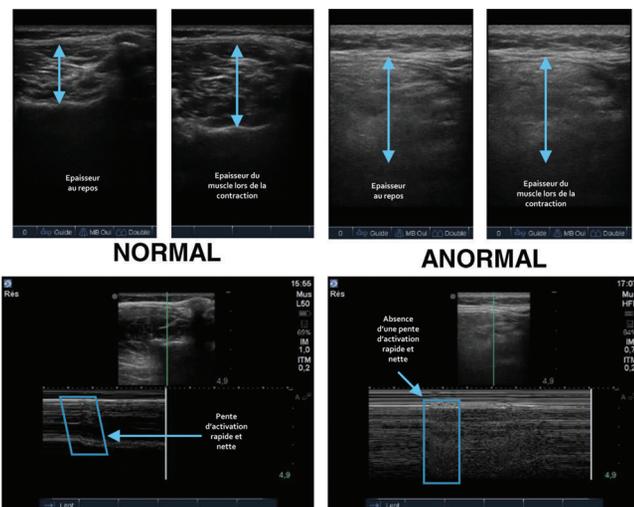


Figure 1. Observation du CRIT18 en mode B et mode TM : pente de contraction nette et rapide des muscles multifides lombaires.



Figure 2. Positions d'évaluations et manœuvres standardisées : ADIM pour le transverse de l'abdomen & élévation du membre inférieur controlatéral pour les multifides lombaires.

un score « global » binaire (présence ou absence de déficit) regroupant les deux côtés (droite et gauche) a été réalisé : si le critère était positif d'un côté au moins, le critère évalué était considéré comme positif (présence d'un déficit).

Le schéma d'activation était analysé qualitativement en mode B et en mode TM. En 2007, Teyhen et al. ont proposé un ensemble de critères d'évaluation du schéma d'activation du muscle TrA (Tableau I). Aucun critère d'évaluation qualitative des MF n'a été retrouvé dans la littérature. Nous avons donc proposé trois critères d'évaluation du schéma d'activation des MF (Tableau I).

L'examineur principal qui réalisait et interprétait les échographies était un kinésithérapeute expérimenté (six années de pratique), formé à l'échographie aux Pays-Bas (3 semaines de formation + 6 mois de supervision clinique) pratiquant l'échographie en cabinet libéral depuis quatre ans (8 à 10 échographies par semaine environ).

Pour étudier la fiabilité et dans le but de minimiser les erreurs liées à la prise de mesures, les images de l'examineur

Tableau I. Critères échographiques qualitatifs pour les muscles multifides et transverse de l'abdomen.

Critères		Interprétation	
Transverse de l'abdomen	Schéma optimal d'activation respecté	1. Raccourcissement du TrA & mise en tension des fascias abdominaux antérieur et thoracolombaire CRIT1	Le TrA doit se raccourcir dans le sens de sa longueur. On doit pouvoir observer un glissement du fascia médial liant le TrA à la ligne blanche et le fascia latéral liant le TrA au fascia thoracolombaire (raccourcissement du TrA et mise en tension)
		2. Le TrA s'épaissit dans sa largeur (contraction) CRIT2	On doit pouvoir observer un épaississement dans le sens de sa largeur lors de la contraction du muscle
		3. le TrA forme un arc latéralement (action de corset) CRIT3	Action de corset. Le muscle s'épaissit et se contracte en forme d'arc (convexe par rapport à l'abdomen) pour contrôler la pression intra-abdominale
		4. Dimensions de l'OE et de l'OI relativement inchangées CRIT4	Pas ou peu de changement d'épaisseurs de l'OE et de l'OI (cela concerne surtout l'OE). L'OI se contracte toujours un peu en fin d'expiration pour y participer
		5. Schéma d'activation symétrique CRIT5	Les critères cités ci-dessus doivent apparaître symétriquement, donc des 2 côtés de manière à peu près égale
	Caractéristiques du schéma global non optimal d'activation	6. Le TrA, l'OI et l'OE s'épaississent tous et augmentent leur largeur simultanément, souvent rapidement CRIT6	Compensation des autres muscles superficiels de l'abdomen avec un épaississement observé
		7. Activation mais le TrA ne se raccourcit pas et n'applique pas de tension sur le fascia adjacent CRIT7	Contraction du TrA sans observer un raccourcissement dans le sens de sa longueur, donc sans mise en tension des fascias adjacents
		8. Le TrA ne s'enveloppe pas autour de la taille CRIT8	Pas d'action de corset, le TrA reste globalement horizontal
		9. Schéma d'activation asymétrique CRIT9	Asymétrie des critères (pas retrouvés des deux côtés)
		10. Maintien de la respiration ou de l'expiration forcée CRIT10	Apnée (contraction des abdominaux superficiels) ou le patient est obligé de forcer l'expiration pour contrôler l'expiration
	Schéma de substitution commun	11. Gainage des muscles abdominaux superficiels CRIT11	Contraction des muscles abdominaux superficiels pour compenser (DA, OE et OI)
		12. Basculement postérieur du pelvis ou flexion du tronc CRIT12	Le ventre se rentre en contractant les abdominaux ou le patient fléchit le tronc ce qui entraîne une contraction des abdominaux superficiels
		13. Dépression de la cage thoracique durant l'ADIM CRIT13	La cage thoracique se rentre en direction du sol (majoration de la contraction des droits de l'abdomen)
		14. Augmentation du poids sur les talons en décubitus CRIT14	Le patient pousse sur ses talons pour augmenter le recrutement musculaire
		15. Activations phasiques rapides et activations contrôlées CRIT15	L'activation est contrôlée dans le sens où l'on n'a pas d'activation progressive et augmentée au fur et à mesure de l'expiration
		16. Peu ou pas de mouvement de la partie basse de l'abdomen CRIT16	Pas de phénomène d'aspiration (<i>drawing-in</i>) de l'abdomen sans contracter les abdominaux superficiels
Multifides	Schéma optimal d'activation respecté	17. Epaissement immédiat CRIT17	L'épaississement doit apparaître tout de suite après avoir donné l'instruction au patient d'élever le membre inférieur (changement d'épaisseur à l'échographie)
		18. Pente de contraction nette et rapide CRIT18	La pente de contraction doit être nette et rapide pour arriver à 100 % de l'épaississement musculaire
		19. Absence de compensation des paraspinaux CRIT19	Les muscles paraspinaux sont observés en mode B et ne doivent pas se sur-activer et entraîner une compression des MF contre le processus épineux et donc diminuer l'épaisseur latérale des MF (absence ou faible épaississement dans cet axe)

TrA : Transverse de l'Abdomen ; OE : Oblique Externe ; OI : Oblique Interne ; DA : Droit Antérieur ; MF : Multifides.



principal étaient réinterprétées, en aveugle, par un second examinateur. Ce dernier était un étudiant en kinésithérapie de dernière année (4 h de formation à l'échographie en formation initiale + 8 échographies réalisées sous la supervision d'un tuteur formé en stage). La fiabilité concernait donc la lecture des images échographiques. Cette option a été préférée car le but était d'obtenir des informations sur l'interprétation des critères. Ainsi, nous avons pu éviter un biais de manipulation par l'expérimentateur junior et limiter la durée des examens échographiques (1 seul examen par patient durant 20 mn maximum).

Analyses statistiques

Fiabilité

Afin de savoir si l'interprétation des images est répétable et reproductible, la fiabilité de chaque critère échographique qualitatif binaire validé dans cette étude a été évaluée (fiabilité inter-examinateur). Le test statistique utilisé est un coefficient kappa de Cohen.

Validité

En l'absence de *gold standard*, on utilise une validité de construit. Le test de référence (les tests de CM) et le test étudié (l'échographie) observent le même paramètre (le contrôle moteur) mais de façon différente. Le test de référence donne un score entre 0 et 6. Si le score est supérieur ou égal à 2, le test est positif pour un déficit de contrôle moteur. Le résultat du test de Luomajoki peut donc être exprimé sous la forme d'une variable qualitative binaire. Le test étudié fournit aussi une variable qualitative binaire (absence de déficit ou présence de déficit) pour chaque critère. L'interprétation échographique se fait, en mode B ou en mode TM, selon le critère évalué. La validité est construite par comparaison des résultats entre les deux évaluations [8]. La concordance réelle a été calculée grâce à la formule du coefficient kappa de Cohen.

Traitement des données

Les données ont été rassemblées dans un tableur puis traitées avec le logiciel RStudio® version 1.1.423 (R version 3.4.3). La fonction "kappa.test" du package "fmsb" a été utilisée pour calculer les coefficients kappa entre les variables qualitatives dans le but de repérer les concordances réelles supérieures à 0,4. Quand le kappa était supérieur ou égal à 0,41, la *p-value* a été calculée. La fonction "PowerBinary" du package "kappaSize" a été utilisée pour calculer le nombre de sujets nécessaires avec une puissance de 80 %. Tous les tests statistiques utilisés étaient des tests bilatéraux avec un risque alpha de 5 %. Afin de limiter le risque de première espèce lié à la multiplicité des tests, la correction de Benjamini-Hochberg a été appliquée sur les *p-values* obtenues (fonction "p.adjust" de R, méthode "BH") [25]. La fonction cor() a été utilisée pour calculer une corrélation simple de Pearson entre le score quantitatif des tests de Luomajoki et le score du Questionnaire Oswestry.

Pour interpréter les valeurs des coefficients kappa, l'échelle de Landis et Koch [8] a été utilisée : "mauvaise" si inférieur à

0,20, "faible" entre 0,21 et 0,40, "moyenne" entre 0,41 et 0,60, "bonne" entre 0,61 et 0,80, "excellente" entre 0,81 et 1,00. Dans cette étude, nous avons décidé de retenir les valeurs kappa supérieures ou égales à 0,41 (moyenne). Ce seuil a été fixé, a priori, lors du calcul du nombre de sujets nécessaires. D'une part, une concordance mauvaise ou faible ne nous semblait pas apporter des informations cliniquement pertinentes. D'autre part un seuil de 0,9 (concordance presque parfaite) souvent recommandé pour valider l'utilisation d'un outil en première intention, nous paraissait trop restrictif pour une étude préliminaire. Afin d'obtenir une puissance de 80 % avec un seuil de 0,41, le nombre de sujets total à inclure était de 50 patients.

RÉSULTATS

Population

L'étude s'est déroulée entre le 1^{er} décembre 2016 et le 1^{er} mai 2017. Le nombre de sujets inclus dans l'étude était de 14 patients lombalgiques dont 5 présentant un déficit de contrôle moteur (groupe « déficit de CM ») selon les tests de Luomajoki (score supérieur ou égal à 2 sur 6) [11] et 9 patients lombalgiques sans déficit de CM lors de ces mêmes tests (groupe « témoin »). Les caractéristiques de l'échantillon sont présentées dans le *Tableau II*. Le recrutement, la randomisation et le suivi de l'ensemble des patients au cours de l'étude sont illustrés dans le diagramme de flux (*Fig. 3*).

Tableau II. Caractéristiques de l'échantillon.

Critères	Groupe déficit de CM	Groupe témoin
	n = 5	n = 9
Homme	0 (0)	3 (33)
Femme	5 (100)	6 (67)
Sportif (> 2 h par semaine)	0 (0)	1 (11)
Âge (Années)		
Global	39,4 [12,3]	41,7 [18,4]
Homme	–	53,3 [7,4]
Femme	39,4 [12,3]	37 [19,2]
Profession		
Travailleur	4 (80)	5 (56)
Sans emploi	1 (20)	0 (0)
Retraité	0 (0)	1 (11)
Étudiant	0 (0)	3 (33)
Score		
Oswestry (score sur 100)	36,8 [26,1]	9,8 [8,9]
Luomajoki (score sur 6)	2,6 [0,89]	0,56 [0,53]
Moyenne [Écart-type] ou Effectif (%).		

Résultats pour les tests

Pour les critères pertinents ($k > 0,4$) les valeurs obtenues lors des tests statistiques (kappa, IC 95 % et p -value) sont présentées dans le *Tableau III*.

Validité de construit

La validité de construit a été étudiée en mesurant la concordance réelle entre chaque critère et le score de Luomajoki (présence d'un déficit de contrôle moteur pour un score supérieur ou égal à 2 sur 6). Lors du calcul des coefficients kappa, trois critères (CRIT15, CRIT16 et CRIT18) affichaient une concordance moyenne avec les tests de Luomajoki (concordance comprise entre 0,41 et 0,60) et un critère (CRIT19) une concordance bonne (comprise entre 0,61 et 0,80). Lors des tests, trois critères étaient statistiquement significatifs (CRIT16, CRIT18 et CRIT19) avec une p -value inférieure à 0,05 après correction de BH.

Fiabilité

La fiabilité inter-examineur a été établie en calculant l'accord (coefficient kappa) entre les deux examinateurs pour les quatre

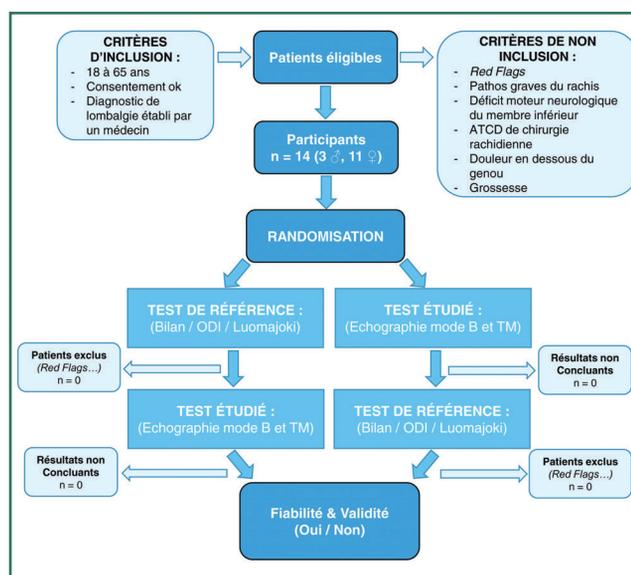


Figure 3. Diagramme de flux de l'étude.

Tableau III. Résultats de l'étude.

Critères	Fiabilité			Validité		
	kappa	p -value	p corrigée ^b	kappa	p -value	p corrigée ^b
CRIT1	n.e	–	–	0,07 (-0,50 ; 0,63)	–	–
CRIT2	n.e	–	–	-0,14 (-0,82 ; 0,55)	–	–
CRIT3	n.e	–	–	0,17 (-0,27 ; 0,60)	–	–
CRIT4	n.e	–	–	-0,24 (-0,72 ; 0,25)	–	–
CRIT5	n.e	–	–	0,31 (-0,17 ; 0,80)	–	–
CRIT6	n.e	–	–	0,04 (-0,46 ; 0,54)	–	–
CRIT7	n.e	–	–	0,07 (-0,50 ; 0,63)	–	–
CRIT8	n.e	–	–	0,26 (-0,18 ; 0,71)	–	–
CRIT9	n.e	–	–	-0,02 (-0,64 ; 0,60)	–	–
CRIT10	n.e	–	–	-0,26 (-0,91 ; 0,40)	–	–
CRIT11	n.e	–	–	0,07 (-0,50 ; 0,63)	–	–
CRIT12	n.e	–	–	-0,26 (-0,91 ; 0,40)	–	–
CRIT13	n.e	–	–	-0,14 (-0,82 ; 0,55)	–	–
CRIT14	n.e	–	–	-0,14 (-0,82 ; 0,55)	–	–
CRIT15	0,57^a (0,14 ; 1)	0,0163	0,0325	0,43 (-0,04 ; 0,90)	0,0540	0,0540
CRIT16	0,47^a (0,03 ; 0,91)	0,0277	0,0370	0,47^a (0,03 ; 0,91)	0,0277	0,0370
CRIT17	n.e	–	–	0,10 (-0,53 ; 0,73)	–	–
CRIT18	0,42 (-0,07 ; 0,90)	0,0633	0,0633	0,55^a (0,11 ; 1)	0,0235	0,0370
CRIT19	0,71^a (0,35 ; 1)	0,0038	0,0151	0,71^a (0,35 ; 1)	0,0038	0,0151

- : p -value pour valeur du coefficient kappa $< 0,4$ (non testé), k (:): intervalle de confiance à 95 % ; n.e : valeur non évaluée.

^aValeur statistiquement significative.

^bCorrection de Bejamini-Hochberg.



critères échographiques ayant une concordance supérieure à 0,4 lors de l'étude de validité (CRIT15, CRIT16, CRIT18 et CRIT19). Elle était moyenne à bonne selon les critères.

Estimations de la précision

Lors de l'étude de fiabilité et de validité, les IC 95 % des coefficients kappa ont été calculés et sont présentés dans le *Tableau III*.

Les scores moyens pour les tests de CM étaient de 2,6 tests positifs (écart-type 0,89) dans le groupe déficit de CM et de 0,56 (écart-type 0,53) dans le groupe contrôle. La corrélation entre les tests de Luomajoki et le questionnaire Oswestry était modérée (cor = 0,54, IC 95 % [0,01–0,83] et $p < 0,05$). Cette corrélation a été calculée à titre indicatif afin d'apporter une information sur un lien potentiel entre déficit de contrôle moteur et incapacité.

La corrélation entre les critères échographiques et une latéralisation potentielle de la lombalgie n'a pas pu être évaluée car l'échantillon était trop restreint.

DISCUSSION

Interprétation des résultats

Cette étude a permis de faire apparaître trois critères qualitatifs valides pour mettre en évidence un déficit de contrôle moteur (CRIT16, CRIT18 et CRIT19).

Lors de l'étude de fiabilité, seuls deux critères (CRIT16 et CRIT19) étaient fiables ($p < 0,05$). Le résultat non significatif pour le troisième critère (CRIT18) pourrait être lié à la mauvaise qualité de certains clichés échographiques, ce qui aurait pu entraîner un biais d'interprétation pour l'examineur junior. Par ailleurs, lors de l'étude de validité, un critère (CRIT15) était à la limite de la significativité statistique (p corrigée = 0,0540). Cette valeur pourrait s'expliquer par un manque de puissance statistique (petit échantillon) lors de notre étude.

Le protocole échographique proposé dans cette étude était réalisé en moins de 20 mn et même en ajoutant à cette durée le temps nécessaire pour interpréter trois critères échographiques qualitatifs, l'examen reste réalisable en 30 mn. Le critère d'applicabilité clinique du protocole était donc respecté. Finalement, l'utilisation d'un examen échographique qualitatif, pour évaluer un déficit de contrôle moteur chez le patient lombalgique non spécifique, semble envisageable au vu des résultats obtenus dans cette étude.

Limites de l'étude

Pour ce qui est de la conception et de l'exécution du protocole d'étude, nous avons relevé quelques pistes d'améliorations possibles.

Tout d'abord, il aurait été préférable que la réalisation et l'interprétation de chaque test aient été réalisées par des examinateurs différents. Idéalement, les clichés échographiques auraient pu être interprétés par un radiologue indépendant

et les tests de Luomajoki par un kinésithérapeute extérieur. Néanmoins, cette façon de procéder aurait nécessité des ressources humaines et financières non disponibles. C'est pourquoi, dans notre étude, même si chaque évaluateur n'avait pas connaissance du résultat de l'autre examen, on ne peut pas exclure la présence d'un biais d'évaluation. Cependant, les valeurs obtenues lors des tests de Luomajoki étaient cohérentes avec les valeurs de références proposées par l'auteur [9] : une moyenne de 2,6 tests positifs dans le groupe déficit de CM contre 2,2 tests positifs dans l'étude de Luomajoki (IC 95 % = 1,94–2,48) et une moyenne de 0,56 test positif dans notre groupe contrôle contre 0,75 test positif dans l'étude de référence (IC 95 % = 0,55–0,95). Le seuil signant un déficit de contrôle moteur étant de 2 sur 6.

Il est possible que la clause d'ignorance (*concealment*) n'ait pas toujours été respectée pour des raisons d'organisation et de logistique inhérentes à l'activité libérale des examinateurs. En effet, les examinateurs géraient eux-mêmes leur planning et avaient la possibilité de connaître l'ordre de passage des tests avant d'inclure les patients. Aussi, la participation d'un patient à l'étude aurait pu être conditionnée par des raisons pratiques ou de disponibilités. Cependant, nous considérons qu'il est peu probable que l'ordre de passage des examens ait pu réellement impacter le résultat des tests. Ces considérations sont toutefois intéressantes dans l'optique de futures études cliniques en cabinet libéral.

La qualité des deux examens utilisés dans cette étude est également très dépendante de la coopération du patient ainsi que de sa capacité à comprendre le test et à le réaliser.

Notons que nous n'avons essuyé aucun refus pour réaliser un test au cours de cette étude. En l'occurrence, la coopération semble donc moins problématique que la réalisation. Aucun patient n'a été exclu pour une incapacité à réaliser les examens demandés. Cependant, il est difficile de savoir dans quelle mesure le fait de ne pas connaître les tests à l'avance a pu influencer les résultats des patients. De la même façon, nous ne disposons d'aucune donnée permettant de connaître le seuil à partir duquel les résultats des patients s'amélioreraient en raison de l'entraînement au test. Le nombre de tentatives avait donc été fixé arbitrairement d'après les données de l'étude de Luomajoki [9].

Pour finir, en ce qui concerne l'étude de fiabilité, l'interprétation des critères échographiques était basée sur une description écrite de ces derniers. Il est intéressant de noter que pour l'examineur novice (seconde interprétation à l'aveugle), il aurait été préférable d'avoir, en plus du texte écrit, des illustrations explicitant chaque critère : deux clichés échographiques de référence à chaque fois, un illustrant le critère dans son état positif et l'autre dans son état négatif. Au vu du faible niveau de formation de l'examineur junior, les résultats de l'étude de fiabilité sont à considérer avec précaution. D'une manière plus générale, la principale limite de notre étude concerne la taille de l'échantillon. Ainsi, les IC 95 % des coefficients kappa présentés dans le *Tableau III* sont très grands pour les valeurs calculées. Ce faisant, cette étude ne permet pas d'estimer avec précision les valeurs de ces coefficients même si nous avons obtenu des résultats

significatifs ($p < 0,05$). Aussi, même quand les valeurs kappa étaient élevées, celles-ci ne peuvent être interprétées autrement que comme étant « mauvaise à excellente ». Par ailleurs, la puissance visée n'a pas été atteinte. Pour être sûr d'obtenir au moins 80 % de puissance, il aurait fallu pouvoir inclure 50 sujets au total. Ce faisant, il est possible que certains critères non significatifs le soient à cause d'un manque de puissance.

Enfin, la validité de construit proposée repose sur l'hypothèse qu'un déficit de contrôle moteur peut être identifié par la batterie de tests de Luomajoki ainsi que par les critères échographiques choisis. Cependant, l'étude de validité des tests de Luomajoki n'a pas été reproduite à ce jour. De la même façon et dans la mesure où la physiopathologie de la lombalgie non spécifique reste imprécise et méconnue, il n'est pas exclu que d'autres critères échographiques pertinents existent et n'aient pas été évalués dans cette étude.

CONCLUSION

Cette étude se positionne comme une étape préalable avant de pouvoir établir la fiabilité et la validité de l'échographie dans l'évaluation du contrôle moteur chez les patients LNS. La fiabilité inter-examineur testée, entre examinateur novice et examinateur sénior, était très variable (moyenne à bonne) selon les critères et devrait maintenant être évaluée entre deux examinateurs confirmés. La validité de l'échographie par rapport aux tests de Luomajoki paraît réelle pour certains critères. En outre, les critères les plus prometteurs pour l'évaluation du contrôle moteur ($k > 0,41$ et $p < 0,05$ pour tous les critères suivants) étaient un phénomène d'aspiration (*drawing-in*) de l'abdomen pour le TrA (CRIT16) ainsi qu'une absence de compensation des muscles paraspinaux pour les MF (CRIT19). Ce sont donc les critères à privilégier lors de l'évaluation et dans des études à venir. Pour finir, l'utilisation de l'échographie en kinésithérapie pour l'évaluation du contrôle moteur semble prometteuse car elle pourrait apporter des informations susceptibles de modifier la prise en soin des patients concernés. En effet, dans l'optique où il existerait au moins un critère fiable et valide par groupe musculaire impliqué dans la lombalgie (transverse, multifides, diaphragme et plancher pelvien) [27], ce type d'évaluation du contrôle moteur permettrait une approche par groupe musculaire là où les tests de Luomajoki permettent une approche par fonction (flexion, extension etc.) [9]. Néanmoins, l'intérêt clinique de ces modifications potentielles reste à établir, et les résultats mis en avant dans cette étude préliminaire devront faire l'objet d'une confirmation par des études de plus grande ampleur.

Contributions

V.L. était investigateur principal de l'étude. Il a écrit le protocole, réalisé l'analyse statistique, analysé les échographies en tant qu'examineur junior. S.M. a réalisé les bilans lombalgiques (tests de Luomajoki et Oswestry). P.P. a supervisé la partie concernant l'analyse statistique. F.Z. a participé à la rédaction et à la relecture de l'article avant publication. A.D. a réalisé et interprété les échographies en tant qu'examineur senior et supervisé le projet.

Déclaration de liens d'intérêts

Anthony Demont est enseignant en formation initiale et continue en échographie musculaire (Société Health Impact). Les autres auteurs déclarent ne pas avoir de liens d'intérêts.

RÉFÉRENCES

- [1] CNOMK. Avis Du Conseil National De L'Ordre Du 27 Mars 2015 Relatif à l'échographie [Internet]. ordremk.fr. 2015.
- [2] Gourmelen J, Chastang JF, Ozguler A, Lanoë JL, Ravaut JF, Leclerc A. Fréquences des lombalgies dans la population française de 30 à 64 ans. Résultats issus de deux enquêtes nationales. *Ann Readapt Med Phys* 2007; 50:640-4.
- [3] Whittaker JL, Teyhen DS, Elliott JM, Cook K, Langevin HM, Dahl HH, et al. Rehabilitative ultrasound imaging: understanding the technology and its applications. *J Orthop Sport Phys Ther* 2007; 37:434-49.
- [4] Teyhen DS, Gill NW, Whittaker JL, Henry SM, Hides JA, Hodges P. Rehabilitative ultrasound imaging of the abdominal muscles. *J Orthop Sports Phys Ther* 2007; 37:450-66.
- [5] Stokes M, Hides J, Elliott J, Kiesel K, Hodges P. Rehabilitative ultrasound imaging of the posterior paraspinal muscles. *J Orthop Sports Phys Ther* 2007; 37:581-95.
- [6] Heidari P, Farahbakhsh F, Rostami M, Noormohammadpour P, Kordi R. The role of ultrasound in diagnosis of the causes of low back pain: a review of the literature. *Asian J Sports Med* 2015;6:e23803.
- [7] Wong AY, Parent EC, Funabashi M, Kawchuk GN. Do changes in transversus abdominis and lumbar multifidus during conservative treatment explain changes in clinical outcomes related to nonspecific low back pain? A systematic review. *J Pain* 2014; 15:377.e1-35.
- [8] Piette P. Métrologie appliquée à la kinésithérapie : mesures, tests et bilans, concepts fondamentaux. EMC - Kine Med Phys Readapt 2016; 0(16):0-15.
- [9] Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Movement control tests of the low back; evaluation of the difference between patients with low back pain and healthy controls. *BMC Musculoskelet Disord* 2008;9:170.
- [10] Delitto A, George SZ, Dillen L Van, Whitman JM, Sowa G, Shekelle P, et al. Low back pain clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. *J Orthop Sport Phys Ther* 2012;42:A1-54.
- [11] Vogler D, Paillex R, Norberg M, de Goumoëns P, Cabri J. Validation transculturelle de l'Oswestry disability index en français. *Ann Readapt Med Phys* 2008; 51:379-85.
- [12] Luomajoki H. Movement control impairment as a sub- group of non-specific low back pain. 2010.
- [13] Luomajoki H, Kool J, de Bruin ED, Airaksinen O. Reliability of movement control tests in the lumbar spine. *BMC Musculoskelet Disord* 2007;8:90.
- [14] Lee DG, Vleeming A. The management of pelvic joint pain and dysfunction. In: Grieve's 3rd Edition.
- [15] Pool-Goudwaard AL, Vleeming A, Stoeckart R, Snijders CJ, Mens JM. Insufficient lumbopelvic stability: a clinical, anatomical and biomechanical approach to « a-specific » low back pain. *Man Ther* 1998; 3: 12-20.
- [16] Richardson C, Jull, Gwendolen, Hodges, Paul. Therapeutic exercise for spinal segmental stabilisation in low back pain. *Spacelist* 1999;(C):125.



- [17] Hodges PW. Is there a role for transversus abdominis in lumbo-pelvic stability? *Man Ther* 1999;4:74-86.
- [18] Rasouli O, Arab AM, Amiri M, Jaberzadeh S. Ultrasound measurement of deep abdominal muscle activity in sitting positions with different stability levels in subjects with and without chronic low back pain. *Man Ther* 2011;16:388-93.
- [19] Jull G a, Richardson C a. Motor control problems in patients with spinal pain: a new direction for therapeutic exercise. *J Manipulative Physiol Ther* 2000; 23:115-7.
- [20] Langevin HM, Stevens-Tuttle D, Fox JR, Badger GJ, Bouffard NA, Krag MH, et al. Ultrasound evidence of altered lumbar connective tissue structure in human subjects with chronic low back pain. *BioMed Cent* 2009; 10:1-9.
- [21] Le Neindre A, Demont A. L'échographie en kinésithérapie: effet de mode ou réel outil d'avenir en rééducation? *Kine Sci* 2015; 43-7.
- [22] Urquhart DM, Hodges PW, Allen TJ, Story IH. Abdominal muscle recruitment during a range of voluntary exercises. *Man Ther* 2005; 10:144-53.
- [23] Teyhen DS, Miltenberger CE, Deiters HM, Toro YM Del, Pulliam JN, Childs JD, et al. The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. *J Orthop Sports Phys Ther* 2005; 35:346-55.
- [24] Miura T, Yamanaka M, Ukishiro K, Tohyama H, Saito H, Samukawa M, et al. Individuals with chronic low back pain do not modulate the level of transversus abdominis muscle contraction across different postures. *Man Ther* 2014; 19:534-40.
- [25] Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *Int J Biostat* 2010; 6: Article 31.
- [26] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Statistical Society Series B (Methodological)* 1995; 57:289-300.
- [27] Lee DG, Vleeming A. The management of pelvic joint pain and dysfunction. In: Grieve's 3rd Edition.