

É

Réadaptation au travail

Études et recherches

RAPPORT R-729



Mesures par ultrasons des muscles profonds du tronc

Évaluation de la fidélité intra et inter-évaluateurs

Programme REPAR - IRSST

*Christian Larivière
Dany Gagnon
Eros de Oliveira Jr.
Sharon Henry
Hakim Mecheri
Jean-Pierre Dumas
Roger Vadeboncoeur*



Solidement implanté au Québec depuis 1980, l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) est un organisme de recherche scientifique reconnu internationalement pour la qualité de ses travaux.

NOS RECHERCHES

travaillent pour vous !

Mission

Contribuer, par la recherche, à la prévention des accidents du travail et des maladies professionnelles ainsi qu'à la réadaptation des travailleurs qui en sont victimes.

Offrir les services de laboratoires et l'expertise nécessaires à l'action du réseau public de prévention en santé et en sécurité du travail.

Assurer la diffusion des connaissances, jouer un rôle de référence scientifique et d'expert.

Doté d'un conseil d'administration paritaire où siègent en nombre égal des représentants des employeurs et des travailleurs, l'IRSST est financé par la Commission de la santé et de la sécurité du travail.

Pour en savoir plus

Visitez notre site Web ! Vous y trouverez une information complète et à jour. De plus, toutes les publications éditées par l'IRSST peuvent être téléchargées gratuitement. www.irsst.qc.ca

Pour connaître l'actualité de la recherche menée ou financée par l'IRSST, abonnez-vous gratuitement au magazine Prévention au travail, publié conjointement par l'Institut et la CSST. Abonnement : 1-877-221-7046

Dépôt légal

Bibliothèque et Archives nationales
2012
ISBN : 978-2-89631-596-3 (PDF)
ISSN : 0820-8395

IRSST - Direction des communications
et de la valorisation de la recherche
505, boul. De Maisonneuve Ouest
Montréal (Québec)
H3A 3C2
Téléphone : 514 288-1551
Télécopieur : 514 288-7636
publications@irsst.qc.ca
www.irsst.qc.ca
© Institut de recherche Robert-Sauvé
en santé et en sécurité du travail,
mars 2012



Réadaptation au travail

Études et recherches

■ RAPPORT R-729

Mesures par ultrasons des muscles profonds du tronc Évaluation de la fidélité intra et inter-évaluateurs Programme REPAR - IRSST

Avis de non-responsabilité

L'IRSST ne donne aucune garantie relative à l'exactitude, la fiabilité ou le caractère exhaustif de l'information contenue dans ce document. En aucun cas l'IRSST ne saurait être tenu responsable pour tout dommage corporel, moral ou matériel résultant de l'utilisation de cette information.

Notez que les contenus des documents sont protégés par les législations canadiennes applicables en matière de propriété intellectuelle.

*Christain Larivière¹, Dany Gagnon², Eros de Oliveira Jr.³,
Sharon Henry⁴, Hakim Mecheri⁵,
Jean-Pierre Dumas⁶, Roger Vadeboncoeur³*

¹Direction scientifique, IRSST

²Université de Montréal

³Institut de réadaptation Gingras-Lindsay de Montréal

⁴Université du Vermont

⁵Prévention des problématiques SST et réadaptation, IRSST

⁶Université de Sherbrooke

Cliquez recherche
www.irsst.qc.ca



Cette publication est disponible
en version PDF
sur le site Web de l'IRSST.

Cette étude a été financée par l'IRSST. Les conclusions et recommandations sont celles des auteurs.

CONFORMÉMENT AUX POLITIQUES DE L'IRSST

Les résultats des travaux de recherche publiés dans ce document
ont fait l'objet d'une évaluation par des pairs.

SOMMAIRE

Introduction. Afin de compléter les outils nécessaires pour évaluer et valider différents programmes d'exercices visant à favoriser le retour au travail chez des travailleurs ayant une lombalgie chronique, des outils d'évaluation de certaines déficiences lombaires doivent être implantés et évalués. Afin d'évaluer les fonctions jouées par les muscles profonds du tronc, des mesures par ultrasons ont été développées. Les objectifs de la présente recherche sont (1) d'importer cette expertise dans notre groupe de recherche (2) de développer les outils d'analyse nécessaires pour extraire les mesures requises des images produites et (3) d'évaluer la fidélité INTRA- et INTER-évaluateurs de ces mesures.

Méthodologie. Quinze sujets sains et 15 sujets lombalgiques chroniques ont été évalués lors de trois séances. Les deux premières séances (en a.m. et p.m.) ont été réalisées avec deux évaluateurs différents, permettant une évaluation de la fidélité INTER-évaluateurs. La troisième séance, a été tenue entre 7 à 14 jours après la dernière séance, réalisée aussi avec l'évaluateur no 1, permettant une évaluation de la fidélité INTRA-évaluateur. Dans chaque séance, quatre tâches standardisées, de faible intensité et permettant de recruter automatiquement certains muscles profonds du tronc, ont été exécutées, deux tâches pour solliciter les multifides lombaires (LM) et deux autres pour solliciter le transverse de l'abdomen (TrA). Celles-ci ont été répétées avec la méthode traditionnelle pour positionner la sonde et avec des méthodes (différentes pour les deux groupes musculaires) ayant le potentiel de contrôler les variations dans le repositionnement de la sonde; ceci, dans le but d'améliorer la fidélité des mesures. Des mesures d'épaisseur des muscles ont été réalisées sur une image prise au repos et une autre prise lorsque la contraction avait été complétée. Une analyse de la fidélité de ces mesures et autres mesures dérivées a été réalisée.

Résultats. Comme attendu, la fidélité INTRA-évaluateur était supérieure à la fidélité INTER-évaluateur, confirmant qu'il est préférable d'utiliser un même évaluateur pour comparer des résultats; ceci, autant pour la recherche que la clinique. Quelques mesures ont démontré une excellente fidélité alors que d'autres ont démontré une fidélité modérée ou pauvre, permettant de sélectionner les mesures les plus prometteuses. Il a été démontré que différentes tâches automatisées peuvent produire des résultats comparables en termes de fidélité, ce qui laisse une plus grande marge de manœuvre aux futurs utilisateurs. Les méthodes évaluées pour mieux contrôler les variations dans le positionnement de la sonde ont permis d'améliorer significativement ($P > 0,05$) la fidélité des mesures, mais ces effets étaient marginaux ($< 2\%$). Des recommandations sur les stratégies à utiliser pour rehausser la fidélité ont été proposées.

Discussion et conclusion. Les mesures obtenues sont comparables à ce qui a été publié ailleurs par d'autres groupes de recherche, démontrant le succès de l'implantation de cette technique de mesure par notre groupe de recherche. Les résultats permettent de se prononcer sur les mesures (variables) à retenir et les mesures à rejeter pour une utilisation future, que ce soit pour des applications en recherche ou en clinique, bien que l'utilité de ce type de mesures des résultats reste à être démontrée avant de justifier un changement de pratique clinique. Bien que les présents résultats soient acceptables pour des applications en recherche, d'autres méthodes doivent être proposées et évaluées pour arriver à mieux standardiser les protocoles d'évaluation et ainsi, réduire les variations inhérentes à ce type de mesures.

REMERCIEMENTS

Nous désirons remercier le Centre de Recherche Interdisciplinaire en Réadaptation du Montréal métropolitain (CRIR – Site Institut de réadaptation Gingras-Lindsay-de-Montréal) pour son soutien en infrastructure de recherche. Le soutien technique a été offert par Daniel Marineau. Soulignons également le suivi administratif du projet, qui a été assuré par Aline Loiseau.

TABLE DES MATIÈRES

1.	INTRODUCTION	1
1.1	Problématique générale des lombalgies chroniques en réadaptation et introduction de la thématique de recherche.....	1
1.2	Objet de l'étude.....	3
1.3	Les déficiences des muscles profonds du tronc et les mesures par ultrasons associées	4
1.4	Fidélité des mesures statiques et dynamiques de l'épaisseur des muscles profonds du tronc	5
1.5	Objectifs de recherche et portée potentielle.....	11
2.	MÉTHODOLOGIE GÉNÉRALE	13
2.1	Objectif 1 : Entraînement de notre équipe de recherche.....	13
2.2	Objectif 2 : Développement d'une interface d'analyse.....	13
2.3	Objectif 3 : Étude de fidélité intra- et INTER-évaluateurs	13
2.3.1	Devis expérimental	13
2.3.2	Sujets.....	13
2.3.3	Évaluateurs.....	15
2.3.4	Procédures.....	15
2.3.5	Tâches	16
2.3.6	Appareil à ultrasons	18
2.3.7	Traitement des images	18
2.3.8	Analyses statistiques	19
3.	RÉSULTATS ET DISCUSSION	21
3.1	Objectif 1 : Entraînement de notre équipe de recherche.....	21
3.2	Objectif 2 : Développement d'une interface d'analyses	22
3.3	Objectif 3 : Étude de fidélité intra- et INTER-évaluateurs	25
3.3.1	Description des données de base.....	25
3.3.2	Étude de la fidélité	32
3.3.2.1	Étude G (composantes de variance).....	32
3.3.2.2	Étude D (stratégies de mesure)	37
3.3.2.2.1	Effet de la tâche et du contrôle du positionnement de la sonde sur la fidélité	39
3.3.2.2.2	Effet de faire la moyenne de trois mesures par image sur la fidélité (muscles abdominaux).....	41
3.3.2.3	Implications pour la recherche et la clinique	42
4.	LIMITES.....	44
5.	CONCLUSION.....	47
6.	IMPACT DU PROJET	49
7.	BIBLIOGRAPHIE.....	51
	ANNEXE A : Résultats de fidélité (Étude D)	58

LISTE DES TABLEAUX

Tableau 1 :	Description des caractéristiques (démographiques, anthropométriques, cliniques) des sujets en fonction du statut (sain, lombalgique) et du sexe	14
Tableau 2 :	Résultats statistiques (probabilités) correspondant aux ANOVAs (2 Contractions × 2 Tâche) appliquées aux différents muscles	31
Tableau 3 :	Pourcentages attribués à chaque composante de variance, selon l'étude G : résultats pour les mesures primaires (au repos et lors de la contraction).....	34
Tableau 4 :	Pourcentages attribués à chaque composante de variance, selon l'étude G : résultats pour les mesures dérivées (DiffRel, CRatio, SRRatio, SCRatio; PTC pour les muscles abdominaux).....	35
Tableau 5 :	Effet de différentes stratégies de mesures (résultats de l'étude D) sur l'amélioration de la fidélité : Pourcentage d'augmentation (ID) ou de diminution (ESM) des indices de fidélité par rapport à la stratégie minimale (référence : 1 image)*	38
Tableau 6 :	Comparaison des indices de fidélité entre la présente étude et les étendues de valeurs rapportées dans la littérature sur le sujet*	40
Tableau 7 :	Résultats statistiques (valeurs <i>P</i>) correspondant aux Test-t (pour données paires) pour comparer les IDs et les ESMs correspondant à différentes tâches et au contrôle du positionnement de la sonde.....	41
Tableau 8 :	Résultats statistiques correspondant à la comparaison entre les trois mesures effectuées pour chaque muscle abdominal, sur chaque image.	42

LISTE DES FIGURES

Figure 1 :	Procédures d'utilisation d'un gabarit pour permettre le repositionnement de la sonde pour les mesures des multifides lombaires (LM)	8
Figure 2 :	Cube de mousse à haute densité (3 × 3 × 3 pouces) dans lequel la sonde est immobilisée à l'aide d'une bande velcro (effet de ceinture autour du cube). Cela permet de diminuer les variations d'inclinaison et de pression sur la sonde.	9
Figure 3 :	Procédures d'utilisation d'un cube de mousse à haute densité (« foam ») pour le contrôle de l'inclinaison et de la pression appliquée sur la sonde pour les mesures du transverse de l'abdomen (TrA).	10
Figure 4 :	Tâches automatisées choisies pour l'étude de fidélité. Les élévations de 5 cm étaient approximative, bien que validées par observation par l'évaluateur, et n'étaient donc pas contrôlées de façon stricte avec une mesure.	17
Figure 5 :	Identification de l'appareil à ultrasons et de ses périphériques (sondes).....	18
Figure 6 :	Illustration des images ultrasons au repos (A) et lors de la contraction (B) du muscle multifide (LM) telles que sélectionnées et visualisées dans l'interface d'analyse Matlab.	23
Figure 7 :	Illustration des images ultrasons au repos (A) et lors de la contraction (B) du muscle transverse de l'abdomen (TrA), oblique interne (OI) et oblique externe (OE) telles que sélectionnées et visualisées dans l'interface d'analyse Matlab. ...	24
Figure 8 :	Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (Repos, Contraction, DiffRel) pour les muscles LM et TrA, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM) et avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le gabarit (pour LM) ou le cube de mousse (pour TrA). Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.	26
Figure 9 :	Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (CRatio, SRRatio, SCRatio) pour les muscles LM et TrA, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM) et avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le gabarit (pour LM) ou le cube de mousse (pour TrA). Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.	27
Figure 10 :	Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (Repos, Contraction, DiffRel) pour les muscles OI et OE, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.	28
Figure 11 :	Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (CRatio, SRRatio, SCRatio) pour les muscles OI et OE, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.	29

- Figure 12 : Statistiques descriptives (moyennes et écarts-types) correspondant à la variable PTC (proportion par rapport épaisseur totale abdominale) pour les trois muscles abdominaux (TrA, OI et OE), aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.30
- Figure 13 : Statistiques descriptives correspondant à la variable DiffRel pour les trois muscles abdominaux (TrA, OI et OE) et les trois tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables (évaluateurs, jours, côtés, images) ont été agrégés pour produire ces statistiques (n = 180). Ces résultats correspondent aux 15 sujets sains seulement.31
- Figure 14 : Illustration des conditions expérimentales qui ont généré des différences significatives (ANOVAs pour mesures répétées; $P < 0,05$) entre les évaluateurs pour le muscle TrA (conditions expérimentales avec le cube de mousse). T1 : ÉAJT; T2 : ÉAJF36
- Figure 15 : Illustration des conditions expérimentales qui ont généré des différences significatives (ANOVAs pour mesures répétées; $P < 0,05$) entre les évaluateurs pour le muscle LM à L4. Les figures du haut et du bas correspondent aux conditions sans (haut) et avec (bas) utilisation d'un gabarit (acétate).36

LISTE DES ABRÉVIATIONS ET SYMBOLES LES PLUS UTILISÉS

Abréviations :

ADIM : « Abdominal drawing-in maneuver »
CCIC : coefficients de corrélation intraclasse
CRatio (Ratio de contraction) : = Contraction / Repos
ÉAJF : élévation active jambe fléchie
ÉAJT : élévation active jambe tendue
ESM : erreur standard de mesure
DiffRel : (Différence relative) = $100 \times (\text{Contraction} - \text{Repos}) / \text{Repos}$
IMC : indice de masse corporelle
ID : index de dépendabilité
L2 ou L4 : 2^e et 4^e vertèbres lombaires
LM : « lumbar multifidus » ou multifides lombaires
LM-bras : tâche de lever du bras (et poids libre) pour solliciter le LM
LM-jambe : tâche de lever de la jambe pour solliciter le LM
OE : oblique externe
OI : oblique interne
PTC (TrA, OE, OI) : proportion du muscle par rapport épaisseur totale abdominaux lors de la contraction
PPDD : plus petite différence détectable
SCRatio (Ratio de symétrie contraction) = Contraction muscle g. / Contraction muscle dr.
SRRatio : (Ratio de symétrie au repos) = Repos muscle gauche / Repos muscle droit
TrA : Transverse de l'abdomen

Symboles :

σ_s^2 : Variance expliquée par la facette Sujet
 σ_E^2 : Variance expliquée par la facette Évaluateur
 σ_J^2 : Variance expliquée par la facette Jour
 σ_I^2 : Variance expliquée par la facette Image
 σ_{SJ}^2 : Variance expliquée par l'interaction Sujet×Jour
 σ_{SI}^2 : Variance expliquée par l'interaction Sujet×Image
 σ_{JI}^2 : Variance expliquée par l'interaction Jour×Image
 σ_{SIJ}^2 : Variance expliquée par l'interaction Sujet×Jour×Image
 σ_{SE}^2 : Variance expliquée par l'interaction Sujet×Évaluateur
 σ_{EI}^2 : Variance expliquée par l'interaction Évaluateur ×Jour
 σ_{SEI}^2 : Variance expliquée par l'interaction Sujet×Évaluateur ×Image

1. INTRODUCTION

1.1 Problématique générale des lombalgies chroniques en réadaptation et introduction de la thématique de recherche

Les maux de dos affectent entre 58 % et 84 % des gens à un moment ou à un autre de leur vie (Dionne 1999), ce qui en fait le plus important problème de santé dans les pays industrialisés (Hildebrandt 1995 ; Spitzer *et coll.* 1987). Uniquement au Québec, la Commission de la Santé et de la Sécurité du Travail (CSST) estime avoir déboursé 516 millions de dollars durant l'année financière 2007 pour les maux de dos (CSST 2008). Toujours selon ce rapport, les maux de dos représentent tout près de 30 % de l'ensemble des lésions professionnelles indemnisées par la CSST et la région lombaire est touchée dans 60 % des cas d'affections vertébrales survenues entre 2004 et 2007. Il devient donc impératif de développer des outils de mesures valides et fidèles permettant de quantifier la présence de déficiences ou d'incapacités associées à la colonne vertébrale lombaire ainsi que des modalités de traitement efficace, afin de favoriser le retour au travail et de réduire les coûts.

Étant donné qu'il n'est pas possible d'attribuer un diagnostic spécifique dans plus de 80 % des cas, ces lombalgies ont le plus souvent un caractère non spécifique. Dans le présent rapport, nous entendrons par lombalgie non spécifique toute lombalgie sans composante neurologique et sans pathologie rachidienne suspectée telle qu'évaluée lors de l'examen clinique aux fins de triage, c'est-à-dire sans l'utilisation de tests para-cliniques (CLIP 2006). L'histoire naturelle des lombalgies est généralement décrite à l'aide d'une classification en trois phases (aiguë, subaiguë, chronique) définies par la persistance des symptômes en relation avec l'absence de travail (Frank *et coll.* 1996 ; Spitzer *et coll.* 1987). Heureusement, la grande majorité des cas (70 %) se règle spontanément durant la phase aiguë, soit dans les 3-4 semaines suivant les premiers symptômes (Frank *et coll.* 1996), bien que les récurrences soient fréquentes. Vingt pour cent additionnels des cas se régleront durant la phase subaiguë, soit dans une fenêtre de 4 à 12 semaines. Malheureusement, une minorité de cas, soit environ 10 % d'entre eux, atteindront la chronicité se définissant par plus de 12 semaines d'absence de travail. Cette minorité aura un taux de retour au travail presque nul dans l'année qui suit et engendra la plus grande proportion des coûts (Hashemi *et coll.* 1997).

Une revue des différents bilans de la littérature conclut que les exercices physiques sont efficaces pour réduire la douleur et les incapacités chez les gens souffrant d'une lombalgie subaiguë ou chronique (Smidt *et coll.* 2005). Cela est d'autant plus vrai lorsque les modalités d'exercice sont actives et ont un dosage suffisant pour produire les effets requis (Hayden *et coll.* 2005 ; Vuori 2001). Ces résultats s'appliquent à trois grandes catégories d'exercices : (1) les **exercices aérobiques** de type *workout* ou sur bicyclette stationnaire, visant principalement l'augmentation de la capacité aérobique (2) les **exercices de renforcement** visant principalement l'augmentation de la force et/ou de l'endurance des muscles du tronc et (3) les **exercices de stabilisation lombaire** visant l'amélioration de la coordination musculaire (contrôle moteur) des muscles profonds du tronc. Jusqu'à présent, la recherche n'a pas réussi à démontrer la supériorité d'une modalité spécifique d'exercice plutôt qu'une autre (CLIP 2006 ; Hayden *et coll.* 2005 ; van Tulder *et coll.* 2000). Ces résultats équivoques peuvent suggérer que les **mécanismes** d'action mis à contribution pour différentes modalités d'exercice sont **d'origine centrale** et agissent via

des médiateurs psychologiques (Smeets *et coll.* 2006). Cependant, cela n'exclut pas l'implication possible de mécanismes d'origine biomécanique (Panjabi 2006) ou neurophysiologiques (Pruimboom and van Dam 2007). En fait, une hypothèse complémentaire serait associée au fait que les patients souffrant d'une lombalgie non spécifique ne représentent pas une population homogène, autant en terme de facteurs psychosociaux que de facteurs physiques ou lésionnels. Conséquemment, il a été proposé que des patients ayant diverses pathologies sous-jacentes répondraient différemment à différentes modalités d'exercices (Leboeuf-Yde and Manniche 2001). Les récentes évidences appuient cette hypothèse en démontrant que différents sous-groupes de patients répondent inégalement à diverses modalités d'exercices (Brennan *et coll.* 2006 ; Childs *et coll.* 2004 ; Hides *et coll.* 2001 ; Long *et coll.* 2004 ; O'Sullivan *et coll.* 1997). Considérant les critères utilisés pour définir les sous-groupes étudiés dans ces études, qui se veulent plus associés à la nature de la lésion sous-jacente, ces résultats impliquent que des **mécanismes périphériques**, donc de nature physique (physiologiques, biomécaniques), puissent aussi contribuer au rétablissement des patients.

L'exercice peut induire des changements physiologiques (force et endurance musculaires, contrôle moteur) favorables au maintien de la **stabilité lombaire** d'un point de vue mécanique, mais a aussi une influence sur des facteurs psychologiques, qui à leur tour peuvent avoir des bienfaits sur la douleur et l'incapacité prolongée. Le présent projet touche les éléments en relation avec la mécanique de la colonne vertébrale lombaire, ce qui nécessite la définition du concept de stabilité lombaire. La stabilité lombaire est un concept ambigu qui s'est mérité plusieurs interprétations (Reeves *et coll.* 2007 ; Eisenstein 1999) et bien que des efforts soient faits pour mieux la décrire (Reeves *et coll.* 2007), aucune définition ne semble faire consensus. Voici donc quelques définitions permettant d'identifier les éléments clés de ce concept :

1. La stabilité mécanique est définie simplement par l'habileté à soutenir des forces additionnelles sans altération importante de la position articulaire (mouvements inter vertébraux : hypermobilité) ou sans dommage des tissus entourant l'articulation.
2. La stabilité mécanique peut aussi être définie par l'habileté d'un système à retourner à sa position initiale suite à une perturbation (définition théorique de la stabilité en mécanique).
3. D'un point de vue clinique, la stabilité peut aussi être vue par l'habileté de la colonne vertébrale à limiter, sous l'effet de chargements physiologiques, les déplacements articulaires de façon à ne pas endommager ou irriter les différents tissus de la colonne (définition par l'hypermobilité).

Aucune mesure acceptable de la stabilité lombaire n'existe en ce moment (Fritz *et coll.* 1998). Les mesures cliniques basées sur l'imagerie (amplitude articulaire) sont insuffisantes, car ne considère pas le comportement dynamique de la colonne lombaire. L'évaluation mécanique de la stabilité articulaire faite à l'examen clinique, de son côté, présente des problèmes de validité et de fidélité. Des mesures indirectes doivent donc être réalisées pour obtenir une appréciation grossière de la stabilité lombaire, ce qui fait l'objet de nos travaux de recherche. Il s'agit des mesures de proprioception lombaire et d'équilibre, des réponses préparatoires et réflexes des muscles du tronc et des mesures ultrasonographiques des muscles profonds. Différentes revues de la littérature font clairement ressortir que les sujets lombalgiques chroniques ont des déficits en ce qui a trait à ces éléments fondamentaux du contrôle de la stabilité lombaire (Ebenbichler *et coll.* 2001 ; Hodges and Moseley 2003 ; Panjabi 2003).

Les évidences actuelles suggèrent que la recherche doit aller au-delà de la comparaison entre différents programmes d'exercices à l'aide d'essais cliniques randomisés appliqués sur des populations hétérogènes de patients. Cette recherche doit plutôt déterminer (1) quels types de patient (profil clinique) répond le mieux à chaque type de programme d'exercices (arrimage patient/intervention) et (2) quels mécanismes d'action sont impliqués lorsque cet arrimage s'avère fructueux. L'arrimage patient/intervention, permettant d'attribuer des modalités d'intervention spécifiques à certains sous-groupes de patients et d'obtenir des résultats cliniques plus importants (Brennan *et coll.* 2006 ; Childs *et coll.* 2004 ; Fritz *et coll.* 2003 ; Long *et coll.* 2004), a aussi le potentiel de réduire les coûts de l'intervention clinique (Childs and Cleland 2006). Pour permettre cet arrimage, un système de classification doit être développé et pour que ce système soit utile sur le plan clinique, il doit idéalement être basé sur l'identification des mécanismes sous-jacents qui entretiennent le problème. D'ailleurs, le développement de méthodes de classification de patients lombalgiques a été identifié comme une priorité de recherche selon plusieurs groupes d'experts (Airaksinen *et coll.* 2004 ; Borkan *et coll.* 1998; Clinical research agenda for physical therapy 2000 ; van Tulder *et coll.* 2000) ainsi que la priorité no 1 des intervenants en première ligne (chiropraticiens, médecins, physiothérapeutes), qui traitent les patients souffrant d'une lombalgie (Henschke *et coll.* 2007). L'étude des mécanismes impliqués lors de différents programmes d'exercices est maintenant une priorité de recherche (Airaksinen *et coll.* 2004 ; Clinical research agenda for physical therapy 2000; Wessels *et coll.* 2006).

Les membres de la présente équipe de recherche s'apprêtent à initier une recherche dans ce domaine, avec un programme d'exercices qui gagne en crédibilité scientifique (Barr *et coll.* 2007 ; McGill 2002 ; Richardson *et coll.* 2004) et en popularité auprès des physiothérapeutes : les exercices de stabilisation lombaire. Pour y arriver, les mesures de résultats permettant d'évaluer les mécanismes pathophysiologiques ciblés par ce programme d'exercices spécifiques ont besoin d'être implantées et évaluées. Dans la présente étude, nous nous attarderons aux mesures par ultrasons permettant d'évaluer la grosseur et le niveau d'activation des muscles profonds des abdominaux et des dorsaux. Ces mesures peuvent aussi être utilisées pour évaluer les effets d'autres programmes d'exercices, plus particulièrement les programmes de renforcement, qui ciblent aussi les muscles du tronc. Jusqu'à maintenant, nous ne savons pas si différents programmes d'exercices ont des effets plus ou moins spécifiques sur les muscles profonds ou plus superficiels du tronc, ce qui pourrait être mesuré avec ces mesures par ultrasons.

1.2 Objet de l'étude

L'objet de l'étude concerne la mesure de l'épaisseur des muscles profonds de l'abdomen et du dos, telle que justifiée et décrite dans la prochaine section. L'objectif principal du projet consiste à importer l'expertise propre à ce type de mesure. Pour ce faire, trois actions ont été posées : (1) Sharon Henry, de l'Université du Vermont, a formé les membres de notre équipe en ce qui a trait aux procédures à suivre pour obtenir des images claires des muscles ciblés et sur la façon de quantifier l'épaisseur de ces muscles à partir de ces images; (2) développement d'une interface d'analyse des images par ultrasons (dans le logiciel Matlab) afin de recueillir les mesures spécifiques à nos applications; (3) réalisation d'une étude de fidélité intra- et INTER-évaluateurs, de manière à démontrer notre capacité à utiliser ce type de mesure.

1.3 Les déficiences des muscles profonds du tronc et les mesures par ultrasons associées

Les évidences s'accumulent en ce qui a trait aux déficiences neuromusculaires observées chez les patients souffrant d'une lombalgie ou chez les personnes ayant déjà eu des épisodes antérieurs. Différentes observations, souvent interreliées, ont été documentées. Premièrement, une atrophie des muscles multifides lombaires (LM), qui est spécifique au côté et au niveau vertébral de la douleur [ipsilatérale aux symptômes, (Barker *et coll.* 2004 ; Hides *et coll.* 1994)] et aussi spécifique relativement aux muscles superficiels (Danneels *et coll.* 2000), a été observée. Deuxièmement, la contraction du LM et du muscle transverse de l'abdomen (TrA), qui normalement se produit de façon anticipée à un mouvement d'un membre, est retardée et diminuée par rapport à des sujets sains (Hodges and Moseley 2003; MacDonald *et coll.* 2009). Ce phénomène est parfois observé de façon concomitante avec l'oblique interne (OI) de l'abdomen (Hodges and Richardson 1999), ce qui n'est pas le cas pour le muscle plus superficiel que constitue l'oblique externe (OE) (Ferreira *et coll.* 2004 ; Hodges and Richardson 1999). Finalement, une diminution du contrôle volontaire du LM et du TrA a aussi été rapportée (Henry and Teyhen 2007). Même si ces déficiences neuromusculaires ont parfois été observées pour les muscles superficiels du tronc, le LM et le TrA ont montré des résultats anormaux de façon plus consistante (Hodges and Moseley 2003), ce qui explique pourquoi beaucoup d'attention a récemment été accordée à ces muscles et pourquoi ils seront aussi ciblés dans la présente étude portant sur les mesures par ultrasons. Il est suggéré que ces déficiences, toutes associées à un manque de contrôle fin de ces muscles, soient ciblées dans un programme d'exercices afin d'améliorer la stabilité lombaire et ainsi diminuer le risque d'épisodes de récurrences et les incapacités (Moseley and Hodges 2005). D'ailleurs, certaines études récentes supportent cette hypothèse à l'aide de l'imagerie par ultrasons utilisée comme mesure de résultat pour quantifier la grosseur et le niveau d'activation du TrA (Ferreira *et coll.* 2010; Vasseljen and Fladmark 2010). Une diminution du niveau d'activation du LM, aussi mesurée avec l'imagerie par ultrasons, a également été associée avec la présence de facteurs prédictifs d'un succès clinique avec un programme d'exercices de stabilisation lombaire tel que déterminé avec une règle de prédiction clinique (Hebert *et coll.* 2010 ; Kiesel *et coll.* 2007b). Malheureusement, ces résultats ne permettent toujours pas de constituer une solide relation de cause à effet entre ces déficiences musculaires et la douleur ou les incapacités. Le développement et l'utilisation des mesures par ultrasons dans d'autres études longitudinales permettront de mieux établir cette relation.

L'imagerie par ultrasons est une méthode sécuritaire, non invasive, et facilement accessible qui peut être utilisée pour caractériser la structure et la fonction des tissus mous. C'est seulement dans les dernières années que l'imagerie quantitative par ultrasons, largement facilitée par les avancements récents dans l'imagerie et les analyses assistées par ordinateur, est devenue d'utilisation plus courante dans les applications en recherche (Whittaker *et coll.* 2007). Deux grandes catégories de mesures sont généralement investiguées : les mesures statiques et dynamiques. Il est important de noter que l'utilisation de ces termes ne fait pas nécessairement référence au changement de longueur musculaire. Les **mesures statiques** révèlent la grosseur (épaisseur) du muscle au repos et sont généralement considérées comme une mesure « proxy » de la force musculaire. Les **mesures dynamiques**, de leurs côtés, résultent du changement d'épaisseur (ou de la section droite) du muscle correspondant à une tâche impliquant un effort musculaire et conséquemment, révèlent si le muscle investigué est utilisé lors de cette tâche. Il

est à noter qu'il est beaucoup plus difficile, sinon impossible (pour les abdominaux), de mesurer la section droite de ces muscles (mesure de surface) avec l'imagerie par ultrasons, ce qui a mené à des mesures d'épaisseur du muscle (mesure linéaire) réalisées dans le plan longitudinal. Les mesures linéaires sont hautement corrélées aux mesures de surface, telles que rapportées dans une revue de la littérature pour le LM (Stokes *et coll.* 2007). Cette démonstration n'a pas été réalisée pour le TrA, ce qui nécessiterait une technique d'imagerie permettant de mesurer la section droite de ce muscle (Teyhen *et coll.* 2007). Ces mesures apparaissent donc appropriées pour identifier les patients ayant des inhibitions réflexes, qui empêcheraient le recrutement des muscles profonds comparativement aux muscles superficiels. Pour ce faire, certaines tâches primaires (les plus faciles sur le plan moteur) sont communément utilisées au début des programmes d'exercices de stabilisation lombaire pour recruter spécifiquement les LM et TrA, tout en évitant l'utilisation des muscles superficiels. Les LM ont généralement été recrutés, à l'aide de la rétroaction par ultrasons, dans la position couchée sur le ventre. Cependant, une autre tâche standardisée a été proposée, à tout le moins pour évaluer ce muscle sans nécessiter l'utilisation de la rétroaction par ultrasons. Toujours couché sur le ventre, cette tâche consiste à lever un petit poids libre de quelques centimètres au-dessus de la table d'examen avec le bras controlatéral (Kiesel *et coll.* 2007a). De son côté, le TrA est sollicité, lors d'une tâche qui consiste à entrer le nombril à l'intérieur et vers le haut de l'abdomen [*abdominal drawing-in maneuver* (ADIM)]. Il a été démontré que la difficulté dans l'habileté à recruter spécifiquement les LM et TrA lors de ces tâches spécifiques a le potentiel d'identifier des sous catégories de patients (Kiesel *et coll.* 2007a) qui pourraient éventuellement répondre différemment à des programmes d'exercices variés. D'ailleurs, une étude récente démontre qu'une diminution de l'activation du LM, mais pas du TrA, est associée à la présence de facteurs prédictifs de succès clinique pour les exercices de stabilisation lombaire (Hebert *et coll.* 2010).

La validité de critère de ces mesures statiques et dynamiques a été documentée pour les LM et TrA à l'aide de l'imagerie par résonance magnétique (Hides *et coll.* 2006 ; Hides *et coll.* 1995) et avec l'électromyographie avec électrodes implantées (Ferreira *et coll.* 2004 ; Kiesel *et coll.* 2007a ; McMeeken *et coll.* 2004). L'avantage de l'imagerie par ultrasons est son accessibilité comparativement aux autres techniques d'imagerie des tissus mous (tomographie axiale, imagerie par résonance magnétique). L'imagerie par ultrasons permet aussi l'évaluation de l'activation de muscles profonds qui ne peuvent être atteints avec l'électromyographie de surface. Cependant, malgré l'intérêt grandissant pour appliquer l'imagerie par ultrasons dans la pratique clinique et en recherche, il reste des questions à répondre en ce qui a trait à la fidélité de ces mesures appliquées aux muscles profonds du tronc. Cette information est requise pour exploiter ces mesures à l'aide de protocoles efficaces (courts, fidèles) dans le but de quantifier les changements neuromusculaires qui peuvent être associés à la présence d'une lombalgie et à un programme d'exercices.

1.4 Fidélité des mesures statiques et dynamiques de l'épaisseur des muscles profonds du tronc

La fidélité des mesures par ultrasons a récemment été revue pour les LM (Stokes *et coll.* 2007; Hebert *et coll.* 2009) et TrA (Teyhen *et coll.* 2007 ; Hebert *et coll.* 2009) et des études additionnelles ont été publiées depuis pour le LM (Koppenhaver *et coll.* 2009; Wallwork *et coll.* 2007) et le TrA (Koppenhaver *et coll.* 2009). Les résultats de fidélité, quantifiés avec les

coefficients de corrélation intraclasse (CCIC), varient entre 0,72 et 0,98 pour le LM, alors qu'ils se situent entre 0,62 et 0,98 pour le TrA, ce qui est généralement acceptable pour une utilisation clinique pour la plupart des études répertoriées. Cependant, les études de fidélité sont très variables en ce qui a trait à plusieurs éléments d'importance : (1) la population de sujets, (2) le type de fidélité (intra- et/ou INTER-évaluateurs), (3) le nombre de sujets, (4) l'indicateur utilisé pour quantifier la fidélité [CCIC, erreur standard de la mesure (ESM)], (5) le temps entre les mesures (fidélité intra-évaluateur), (6) l'utilisation ou non d'une procédure pour repositionner la sonde de façon reproductible, (7) l'estimation ou non du nombre d'essais, dont on doit faire la moyenne pour atteindre un niveau de fidélité acceptable. Voici des éléments qui n'ont pas été bien couverts dans les études de fidélité réalisées jusqu'à maintenant, et qui représentent des besoins de recherche :

- De toutes les études de fidélité ($n = 26$), seulement trois ont été menées sur la **population d'intérêt** chez qui ces mesures seront utilisées, soit les patients souffrant d'une lombalgie. Recruter des patients, qui ont le potentiel d'avoir une asymétrie entre les côtés pour le LM, c.-à-d. ceux qui ont une douleur unilatérale (Hides *et coll.* 1994), permettrait d'évaluer la fidélité des variables correspondantes (asymétrie gauche/droite), tout en bénéficiant d'une bonne estimation de la variabilité intersujet, ce qui est important dans le calcul des CCIC;
- Aucune étude n'a rapporté la fidélité de **variables dérivées des mesures individuelles** obtenues sur différents muscles (ex. : asymétrie gauche/droite). Par exemple, il a été proposé de rapporter l'épaisseur des muscles abdominaux en proportion de l'épaisseur totale de tous les muscles abdominaux (TrA + OI + OE), ce qui représente une sorte de normalisation permettant les comparaisons interindividuelles (Rankin *et coll.* 2006). Cependant, bien que ces mesures soient plus utiles sur le plan clinique, elles incorporent les sources de variations (ou erreurs) provenant de toutes les mesures individuelles utilisées, ce qui demande à être examiné. Une autre mesure « dérivée » souffrant de ce problème (Koppenhaver *et coll.* 2009) est le pourcentage de changement entre le muscle au repos et lors de la contraction, ce qui n'a pas non plus souvent été évalué tel que décrit dans le prochain point;
- Peu d'études ont évalué la **fidélité des mesures dynamiques**. Les mesures dynamiques introduisent des sources additionnelles de variations, notamment la contraction en elle-même, qui doit idéalement être encadrée avec une tâche standardisée et des instructions précises. D'ailleurs, pour ce qui est du TrA, la plupart des études ont utilisé la tâche (ADIM) servant à entraîner les patients au début d'un programme d'exercices de stabilisation lombaire. Cependant, l'utilisation de cette tâche comporte deux limites très importantes : (1) l'utilisation d'une contraction volontaire (ADIM) est susceptible de mener à des résultats plus variables qu'une contraction automatisée (ex. : élévation active de la jambe tendue - ÉAJT); (2) une mesure de résultats (« outcome measure ») doit être indépendante du traitement. Pour illustrer ce dernier point, prenons comme exemple l'évaluation de l'efficacité de deux programmes d'exercices pour aider à mieux utiliser le TrA, soit un programme de renforcement et un programme de stabilisation lombaire avec accent sur l'ADIM. Si la mesure de résultats est basée sur l'ADIM, les résultats risquent d'être biaisés, car il est fort possible que les patients, ayant pratiqué l'ADIM tout au long de la réadaptation, soient plus en mesure de démontrer un effet lors de l'évaluation finale, aussi réalisée avec cette même tâche. Une contraction automatisée ne faisant pas partie du programme d'exercices n'introduira pas ce type de biais. Ces éléments ont été pris en compte dans une étude (Koppenhaver *et coll.* 2009), qui a paru au cours du présent projet de recherche.

Étonnamment, la fidélité de l'ADIM était généralement comparable à l'ÉAJT (bien que moins bonne dans certains cas de figure), ce qui pouvait s'expliquer par un changement d'épaisseur moins important lors de l'ÉAJT (0,5 mm au lieu de 2,5 mm pour l'ADIM). Les auteurs recommandent que différentes contractions automatisées soient étudiées, afin de trouver le meilleur protocole pour diminuer les erreurs de mesure;

- L'importance de **repositionner la sonde** (par rapport aux muscles) de manière reproductible entre les mesures répétées et, plus particulièrement, entre des mesures réalisées sur des jours différents, n'a jamais été étudiée. Ce besoin de recherche a d'ailleurs été relevé récemment (Koppenhaver *et coll.* 2009). L'utilisation d'un gabarit permettant de repositionner la sonde au même endroit sur la peau est une technique parfois utilisée pour le LM (Stokes *et coll.* 2005) alors que pour le TrA, l'orientation et la pression de la sonde sont contrôlées en l'introduisant dans un cube de mousse (« foam ») à haute densité (Ferreira *et coll.* 2004). Cependant, nous ne savons pas si ces techniques améliorent la fidélité des mesures. **Le gabarit** est un transparent ou un acétate sur lequel des repères anatomiques facilement palpables et/ou des taches de naissance ou des taches de rousseur sont utilisés pour repositionner la sonde au même endroit sur la peau d'un jour à l'autre (procédures détaillées à la **Figure 1**). Le cube de mousse laisse la tête de la sonde dépasser, de manière à ce que la même pression soit appliquée sur l'abdomen d'une mesure à l'autre (**Figure 2**). Le cube de mousse doit avoir une dimension suffisante pour que sa surface se colle parallèlement à la peau, ce qui assure que l'inclinaison de la sonde reste la même d'une mesure à l'autre. Cependant, le cube de mousse n'est pas applicable pour le LM à cause des formes irrégulières près de la colonne vertébrale et parce qu'il faut toujours varier un peu l'inclinaison de la sonde pour bien voir les apophyses transverses nécessaires à la quantification du LM. De la même manière, il n'a pas été jugé nécessaire d'utiliser un gabarit pour repositionner la sonde au même endroit pour le TrA, car une méthode alternative a été utilisée, basée sur un repère anatomique situé à l'interne (fin du muscle TrA), tel qu'identifié avec la sonde (procédures détaillées à la **Figure 3**).


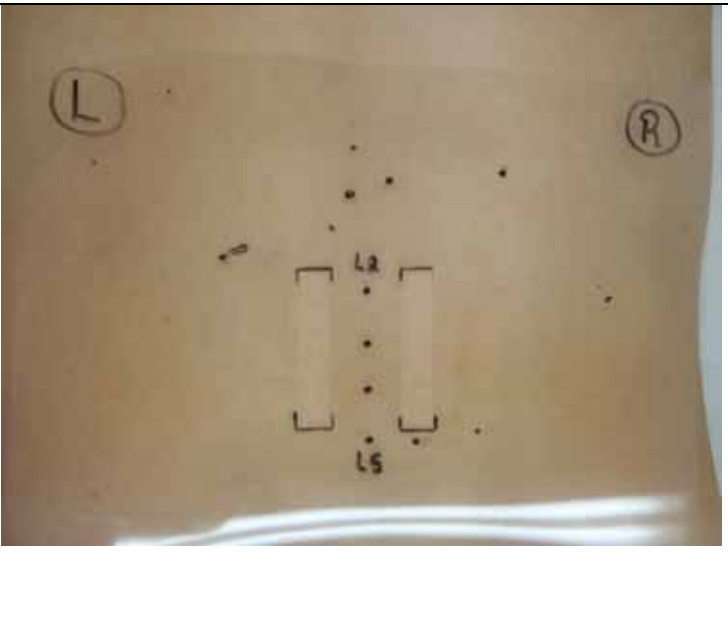
<p>A) Les apophyses épineuses sont d'abord trouvées par palpation et identifiées au crayon. La sonde est ensuite utilisée pour confirmer la position vis-à-vis les bonnes vertèbres en partant du sacrum et en comptant les facettes vers L2. La sonde est ensuite positionnée à la hauteur de L3, à environ 2 cm de la colonne et de manière à bien voir les apophyses transverses. Une fois que l'image est satisfaisante, le contour de la sonde (aux extrémités) est fait avec le crayon.</p>	
<p>B) Les marques faites en A sont calquées sur un acétate placé sur le dos. Ensuite, les différentes taches permanentes (taches de naissance ou de rousseur) sont identifiées sur l'acétate. Ces taches permettront de repositionner l'acétate à la bonne place, lors d'une autre journée d'évaluation (c.-à-d. lorsque les marques faites en A auront été effacées). Finalement, des trous sont faits au ciseau afin de reproduire le contour de la sonde. Lors de la seconde évaluation, ces trous sont marqués sur la peau afin de servir de guide pour positionner la sonde.</p>	

Figure 1 : Procédures d'utilisation d'un gabarit pour permettre le repositionnement de la sonde pour les mesures des multifides lombaires (LM)



Figure 2 : Cube de mousse à haute densité (3 × 3 × 3 pouces) dans lequel la sonde est immobilisée à l'aide d'une bande velcro (effet de ceinture autour du cube). Cela permet de diminuer les variations d'inclinaison et de pression sur la sonde.

- Certains auteurs font trois mesures juxtaposées des muscles abdominaux, à 1 cm d'intervalle (voir Figure 7), car ces muscles ont l'avantage d'avoir une épaisseur relativement constante (Ferreira *et coll.* 2004). Bien que ce ne soit pas explicitement dit, le but est de rehausser la fidélité des mesures, bien que l'effet sur la fidélité n'ait jamais été documenté;
- L'effet des différentes sources de variations a été évalué un à la fois dans cette littérature spécifique (mesures ultrasons des muscles profonds du tronc). Le fait de ne pas **considérer toutes les sources de variations** associées à ce type de mesure dans la même étude de fidélité et en utilisant le même modèle statistique ne permet pas de quantifier leur importance relative. La théorie générale de la fidélité (« generalizability theory ») permet de résoudre ce problème (Shavelson and Webb 1991). En effet, comparativement à l'utilisation classique des CCIC et ESM, cette méthode permet d'évaluer l'efficacité de différents scénarios de mesure afin de trouver un protocole menant à des mesures fidèles tout en demeurant réaliste pour la pratique clinique ou pour la recherche. Jusqu'à maintenant, les sources de variations qui ont été évaluées étaient associées avec : (1) l'évaluateur (fidélité inter-évaluateur), (2) les images ou essais faits au cours d'une même session de mesures (fidélité intra-évaluateur), (3) les lectures d'une même image et (4) les jours. Cependant, il est bien connu que les lectures répétées d'une même image ne mènent qu'à des erreurs négligeables (Hides *et coll.* 2007 ; Teyhen *et coll.* 2005). Il faut donc considérer d'autres stratégies pour améliorer la fidélité des mesures. Nous croyons que le repositionnement de la sonde est une source de variations non négligeable sur laquelle des actions peuvent être posées (gabarit pour le LM, cube de mousse pour le TrA), ce qui constituera une variable indépendante dans la présente étude. Différentes tâches peuvent aussi mener à différents résultats, ce qui sera aussi investigué, bien que cette variable ne sera pas utilisée dans le même modèle statistique.

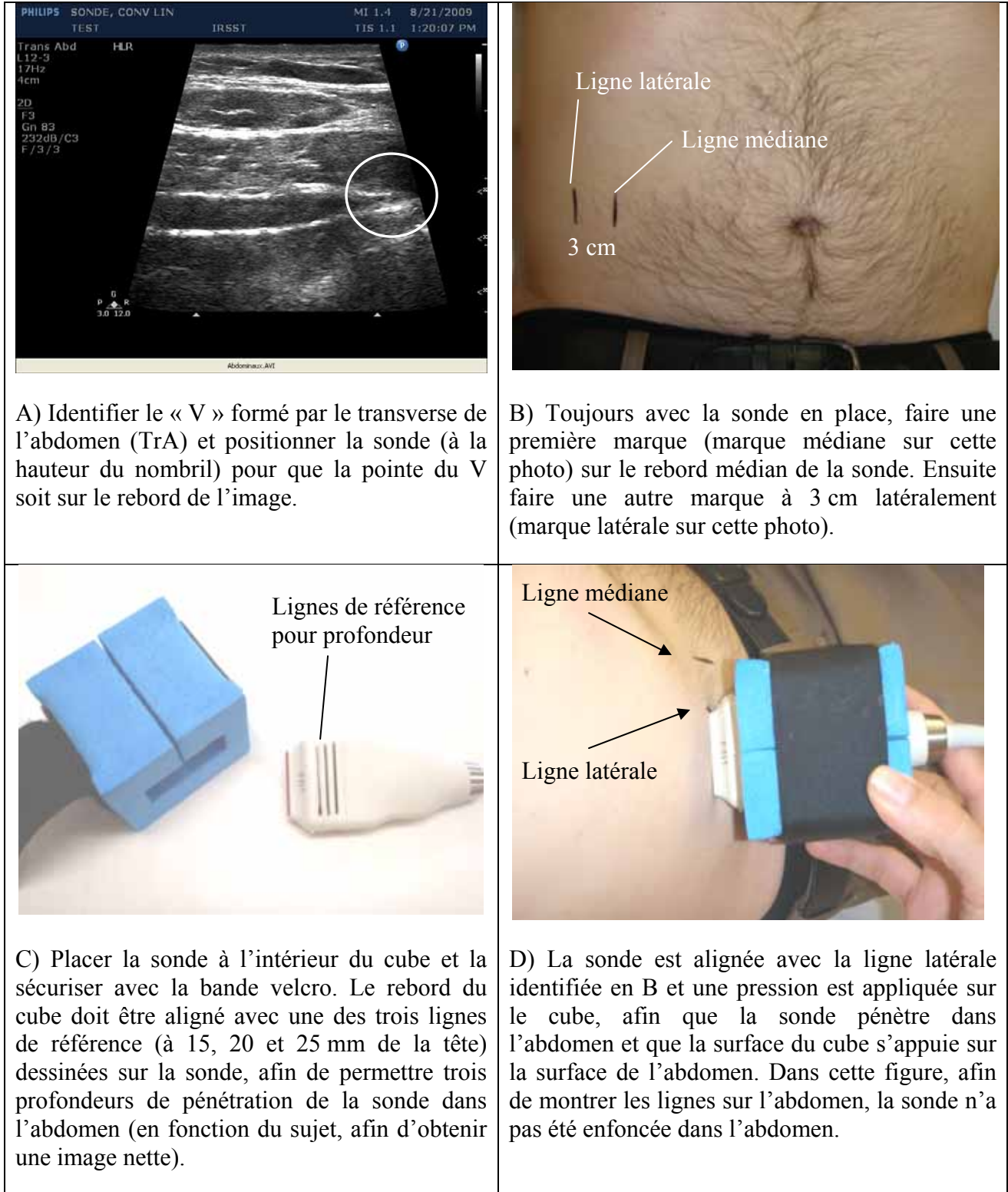


Figure 3 : Procédures d’utilisation d’un cube de mousse à haute densité (« foam ») pour le contrôle de l’inclinaison et de la pression appliquée sur la sonde pour les mesures du transverse de l’abdomen (TrA).

1.5 Objectifs de recherche et portée potentielle

L'objectif principal du projet consiste à s'approprier, selon les meilleures pratiques, l'expertise spécifique connue des mesures par ultrasons des muscles LM et TrA. Pour ce faire, trois objectifs ont été poursuivis :

- (1) Former les membres de l'équipe en ce qui a trait aux procédures à suivre pour obtenir des images claires des muscles ciblés et sur la façon de quantifier l'épaisseur de ces muscles à partir de ces images;
- (2) Développer une interface d'analyse des images par ultrasons (dans le logiciel Matlab), afin de recueillir les mesures spécifiques à nos applications;
- (3) Réaliser une étude de fidélité intra- et INTER-évaluateurs, chez un groupe de sujets sains et un autre groupe de sujets ayant une lombalgie chronique non spécifique, de manière à démontrer notre capacité à utiliser ce type de mesure. Cette étude de fidélité permettra de quantifier l'importance de différentes sources de variations et de déterminer combien d'images (essais) devront être recueillies pour obtenir une fidélité acceptable ($CCIC \geq 0,75$; $ESM \leq 10 \%$). Cette étude permettra aussi d'évaluer différentes possibilités d'améliorer la fidélité, soit : (1) en utilisant des tâches automatisées, (2) en contrôlant le positionnement de la sonde et (3) en effectuant trois mesures juxtaposées (pour les muscles abdominaux).

La portée potentielle de cette étude est de proposer un protocole menant à des mesures fidèles permettant le suivi des patients lors du processus de réadaptation dans de futurs projets de recherche. Cette recherche permettra ensuite de déterminer si ces mesures sont réellement en lien avec la condition des patients. Dans l'affirmative, l'utilisation de ces mesures en milieu clinique pourra être envisagée avec plus de sérieux, ce qui nécessitera la formation des cliniciens (transfert des connaissances) et une législation plus claire dans un contexte de gestes de nature diagnostique.

2. MÉTHODOLOGIE GÉNÉRALE

2.1 Objectif 1 : Entraînement de notre équipe de recherche

Sharon Henry, de l'Université du Vermont et experte dans ce domaine de recherche (Henry and Westervelt 2005 ; Henry and Teyhen 2007), est venue à Montréal pour donner un premier trois heures de formation de manière à : (1) nous aider à reconnaître les structures anatomiques de référence pour positionner la sonde, (2) identifier les muscles d'intérêt, (3) modifier les contrastes pour obtenir des images les plus nettes possibles. Après une période d'environ quatre mois où plusieurs sujets pilotes (n = 9) ont été évalués pour pratiquer les techniques et monter le protocole expérimental, Dr Henry est revenu pour vérifier la capacité de nos évaluateurs à réaliser les mesures demandées (2 heures). Outre cette formation, les activités préparatoires, réalisées lors de la période de quatre mois séparant les visites du Dr Henry, sont détaillées dans la section résultats.

2.2 Objectif 2 : Développement d'une interface d'analyse

Une interface utilisateur a été développée dans le logiciel Matlab par un ingénieur/programmeur. Plusieurs versions ont été utilisées et modifiées pour arriver à un produit final. Il est possible d'en faire un fichier exécutable sur n'importe quel ordinateur, et cela sans nécessiter une licence de Matlab sur le poste utilisé. Un tutoriel (fichier Word) ainsi que des fichiers vidéo ont été constitués pour comprendre le fonctionnement de l'application. L'anglais devait être utilisé pour tous ces livrables, afin de les partager avec nos collaborateurs anglophones.

2.3 Objectif 3 : Étude de fidélité intra- et INTER-évaluateurs

2.3.1 Devis expérimental

Il s'agit d'une étude transversale avec un devis à mesures répétées. Les participants devaient être évalués lors de trois sessions de mesures, les deux premières lors du jour 1 (en a.m. et en p.m.) et la troisième, entre 7 et 14 jours après. Les deux évaluateurs, dont l'ordre a été contrebalancé à travers les sujets, ont été impliqués lors du jour 1 (fidélité INTER-évaluateurs), alors que seul l'évaluateur 1 répéta l'évaluation lors du jour 2 (fidélité intra-évaluateur entre jours 1 et 2). La toute première session de mesures était plus longue (2 h) que les autres (1 h 30), car des questionnaires étaient remplis et parce que le gabarit devait être fait. Toutes les évaluations ont été réalisées au centre de recherche de l'Institut de réadaptation Gingras-Lindsay de Montréal entre novembre 2009 et juillet 2010.

2.3.2 Sujets

Quinze sujets sains et quinze sujets lombalgiques chroniques, répartis également en fonction du sexe, ont été recrutés sur une base volontaire (voir description au **Tableau 1**). Les sujets lombalgiques ont été sélectionnés dans la phase chronique aux fins de l'étude de fidélité (stabilité statut clinique). Les sujets sains ont été recrutés par le bouche à oreille, alors que les sujets lombalgiques étaient recrutés dans la population (annonce dans le journal) et à l'Institut de réadaptation Gingras-Lindsay de Montréal. Les **critères d'inclusion** étaient les suivants : (1) être

âgé entre 18 et 55 ans et (2) maîtriser la langue française ou anglaise. Pour les sujets lombalgiques, les critères d'inclusion additionnels suivants ont été appliqués : (1) la présence de douleur lombaire ou lombo-sacrée unilatérale (pour réaliser une étude comparative; ne représente pas un objectif de la présente étude) avec ou sans douleur radiculaire (limitée aux genoux, afin de correspondre aux patients pour qui les exercices physiques ne sont pas proscrits); (2) la présence de douleur chronique définie comme une douleur quotidienne ou quasi quotidienne depuis plus de 3 mois. Les **critères d'exclusion** étaient les suivants : (1) chirurgie au niveau de l'appareil musculosquelettique du tronc, (2) scoliose ou malformation congénitale de la colonne vertébrale (spondylolyse, fusions intervertébrales, 4 vertèbres lombaires) tel que rapporté par le sujet (pas d'examen radiologique), (3) maladie systémique ou dégénérative (diabète, asthme, arthrite, rhumatisme, sida, cancer, sclérose en plaques ou autre), (4) indice de masse corporelle (IMC) supérieur à 30 kg/m² (Garrow and Webster 1985), (5) réponse positive au Questionnaire d'Aptitude à l'Activité Physique de manière à éliminer les personnes susceptibles de présenter des troubles cardiaques, (6) antécédents d'atteinte neurologique autre que celles découlant des maux de dos (troubles d'équilibre, histoire d'accident vasculaire cérébral), (7) signes neurologiques (réflexes, myotomes, dermatomes) et (8) être enceinte. Un **critère d'exclusion** additionnel a été appliqué chez les sujets sains, soit la présence d'un mal de dos dans la dernière année ou d'un mal de dos d'une durée de plus d'une semaine dans les années précédentes.

Les collectes de données se sont réalisées à l'Institut de réadaptation de Montréal, un établissement du CRIR (Centre de Recherche Interdisciplinaire en Réadaptation (CRIR) du Montréal métropolitain). Le projet a été évalué et accepté par le comité d'éthique du CRIR.

Tableau 1 : Description des caractéristiques (démographiques, anthropométriques, cliniques) des sujets en fonction du statut (sain, lombalgique) et du sexe

Variable *	Sujets sains		Sujets lombalgiques		Valeur P *	
	Hommes (n = 8)	Femmes (n = 7)	Hommes (n = 5)	Femmes (n = 10)	Statut	Sexe
Âge (années)	38 (8)	40 (5)	52 (6)	47 (6)	0,003	0,519
Taille (m)	1,77 (0,04)	1,62 (0,08)	1,73 (0,04)	1,62 (0,04)	0,304	0,000
Masse (kg)	84 (12)	66 (8)	72 (10)	58 (6)	0,006	0,000
IMC (kg/m ²)	26 (3)	25 (3)	24 (3)	22 (2)	0,010	0,091
%graisse	24 (7)	33 (6)	24 (6)	31 (6)	0,606	0,002
Pli cutané LM (mm)	21 (6)	23 (9)	20 (8)	16 (7)	0,164	0,689
Pli cutané TrA (mm)	21 (9)	17 (7)	17 (6)	12 (7)	0,173	0,122
Activité physique	8,2 (1,4)	8,5 (1,1)	7,4 (0,6)	9,3 (1,1)	0,934	0,017
Durée douleur (années)	/	/	16 (12)	12 (11)	/	0,564

IMC : Indice de masse corporelle; % graisse : estimé à partir des plis cutanés (biceps, triceps, abdominal, sous-scapulaire) (Durnin and Womersley 1974); Activité physique telle que mesurée avec le questionnaire de Baecke (Baecke *et coll.* 1982); *Les différences significatives (valeurs

P) sont identifiées en caractères gras. Les interactions Statut × Sexe étaient toutes non significatives.

2.3.3 Évaluateurs

Deux physiothérapeutes (PT1 et PT2) ont recueilli les images alors que pour leur analyse, un des deux a dû être remplacé (identifié comme l'évaluateur no 2 ici) par une kinésiologue, qui a été formée spécifiquement pour l'analyse des images. Ce remplacement a été nécessaire, car le PT2 ne disposait plus d'assez de temps pour analyser lui-même les images, une tâche qui s'est avérée assez exigeante en temps. Cependant, étant donné qu'il n'y a pas de relation entre la prise et l'analyse d'image, ceci ne constitue pas un problème méthodologique. En effet, les habiletés nécessaires pour capturer les images ne sont pas les mêmes que les habiletés pour réaliser les mesures d'épaisseur des muscles.

PT2 était le seul à avoir de l'expérience avec les mesures ultrasons, expérience acquise lors d'une formation postdoctorale d'un an, mais spécifique à une autre application que les muscles du tronc (tendon du biceps). Les PT1 et PT2 ont été formés par Sharon Henry de l'Université du Vermont, une chercheuse reconnue dans ce domaine (Henry and Westervelt 2005 ; Henry and Teyhen 2007 ; Langevin *et coll.* 2009 ; Teyhen *et coll.* 2007). La formation a été constituée des étapes suivantes : (1) une première session de 3 heures avec Dr Henry pour apprendre les techniques, (2) une période de quatre mois de pratique avec des sujets pilotes (n = 9), totalisant 60 heures pour chaque évaluateur et (3) une session de 2 heures avec Dr Henry, afin de valider le protocole de mesures et les techniques utilisées par les évaluateurs. Nous considérons qu'il s'agit d'une formation minimale qui peut correspondre à la formation de nouveaux investigateurs ou cliniciens dans ce domaine d'application.

En ce qui a trait à l'analyse des images, l'expérience acquise par l'évaluateur no 1, qui a participé aux échanges sur les façons de quantifier les images (voir la section résultats – objectif 1) a été transmise à l'évaluateur no 2 (3 heures). Finalement, une étude pilote a été réalisée, afin de comparer les mesures obtenues à partir des mêmes images provenant de quatre sujets. Les écarts de mesures entre les évaluateurs paraissaient acceptables comparativement aux erreurs de mesures déjà disponibles dans la littérature.

2.3.4 Procédures

Les étapes de la cueillette de données ont été les suivantes : (1) formulaire de consentement, (2) questionnaires et mesures anthropométriques; (3) marquage de la peau pour le positionnement de la sonde (faire le gabarit lors de la première session d'évaluation ou utilisation de celui-ci lors des autres sessions); (4) familiarisation (n = 1-2-3 essais) avec chaque tâche (session 1 seulement), (5) prise des images lors des différentes tâches; (6) intensité de la douleur avec l'échelle numérique. Les tâches (n = 4) illustrées à la Figure 4 étaient présentées de la façon suivante. L'ordre des évaluateurs (sessions 1 et 2) a été contrebalancé. L'ordre des muscles (LM ou TrA) était aussi contrebalancé, mais pour chaque muscle, les essais sans accessoire (gabarit/cube) pour repositionner la sonde étaient faits avant les essais avec l'utilisation du gabarit/cube. En ce qui a trait à l'utilisation du cube de mousse, la ligne de référence (à 15, 20 ou 25 mm de la tête; voir Figure 3C) permettant d'indiquer le niveau de pénétration de la sonde dans l'abdomen (contrôle de la pression de la sonde sur la peau) a été notée lors de la première

session de mesure et réutilisée lors des deux autres sessions de mesure. Dans chacune de ces conditions (avec ou sans gabarit/cube), il y avait contre balancement des conditions expérimentales en ce qui a trait au côté du corps ($n = 2$) et à la tâche automatisée ($n = 2$). Pour chaque condition expérimentale, deux essais étaient réalisés un à la suite de l'autre, en retirant la sonde de la peau du sujet avant chaque essai.

2.3.5 Tâches

Les quatre tâches principales sont décrites à la Figure 4. Leur justification est donnée dans les résultats propres à l'objectif 1. À des fins de comparaison avec les tâches ÉAJT et EAJF en terme d'amplitude de contraction musculaire du TrA, une tâche additionnelle, l'ADIM (deux essais, avec le cube seulement, des deux côtés du corps), a été réalisée chez les sujets sains seulement, à la suite des deux tâches pour évaluer le TrA (ÉAJT et élévation active de la jambe fléchie ou ÉAJF). L'ADIM n'a pas été réalisée chez les sujets lombalgiques, car il a été jugé que cela allongait trop la collecte de données, risquant de rendre l'évaluation inacceptable pour certains d'entre eux. Cette tâche est effectuée dans la même position que l'ÉAJF. L'apprentissage de cette tâche était minimal (3-5 essais; requérait environ 2 min) et ne se compare donc pas à la phase initiale d'apprentissage de l'ADIM dans un programme d'exercices de stabilisation lombaire. L'ADIM consistait à rentrer le nombril vers la colonne vertébrale et de le soulever. Une rétroaction avec l'imagerie par ultrasons était donnée afin d'accélérer cet apprentissage. Cette rétroaction consistait à montrer le V du TrA et de demander au sujet de faire bouger ce V latéralement grâce à l'ADIM. Cependant, aucune rétroaction n'était donnée lors des évaluations.

Quelques éléments additionnels doivent s'ajouter à la description de la Figure 4. La tâche LM-bras doit se réaliser avec la tête appuyée sur la table d'examen. Pour les tâches LM-bras et LM-jambe, un ou deux oreillers étaient placés sous l'abdomen, de manière à placer à l'horizontale l'articulation lombo-sacrée, ce qui devait être confirmé par un inclinomètre placé sur le sacrum (angle $\leq 10^\circ$). De manière à minimiser les effets de la respiration sur les mesures, chaque tâche était réalisée une fois qu'une expiration avait été complétée. Le sujet devait donc compléter la tâche avant d'inspirer de nouveau. L'enregistrement (en mode vidéo) durait 7 sec, ce qui laissait au sujet le temps de réaliser la tâche progressivement, c.-à-d. sans gestes brusques, et de tenir la position pendant quelques secondes de manière à stabiliser la contraction musculaire.



A) Tâche LM-bras (Kiesel *et coll.* 2007a): Le **bras controlatéral** (coude à 90°) lève de 5 cm un poids positionné devant la tête. Poids déterminé en fonction de la masse du sujet : (masse < 150 lb : poids de 1,5 lb; masse entre 150 et 175 lb : poids de 2 lb; masse entre 175 et 200 lb : poids de 2,5 lb; masse > 200 lb : poids de 3 lb)



B) Tâche ÉAJT (Teyhen *et coll.* 2007): Élévation active jambe tendue (**jambe controlatérale** au muscle évalué) de 5 cm au-dessus de la table d'examen.



C) Tâche LM-jambe : En position couchée sur le ventre avec les genoux à 90°. Lever de 5 cm la **jambe controlatérale** au muscle évalué.



D) Tâche ÉAJF : En position couchée sur le dos avec les genoux à 90° (« hooklying ») et la hanche à 135°. Lever de 5 cm les **deux jambes** à la fois.

Figure 4 : Tâches automatisées choisies pour l'étude de fidélité. Les élévations de 5 cm étaient approximative, bien que validées par observation par l'évaluateur, et n'étaient donc pas contrôlées de façon stricte avec une mesure.

2.3.6 Appareil à ultrasons

Les équipements, tels que décrits à la Figure 5, proviennent de la compagnie Philips (Philips Medical Systems, Bothell, WA). Les paramètres de l'appareil ont été optimisés pour chaque muscle et réutilisés pour tous les sujets.

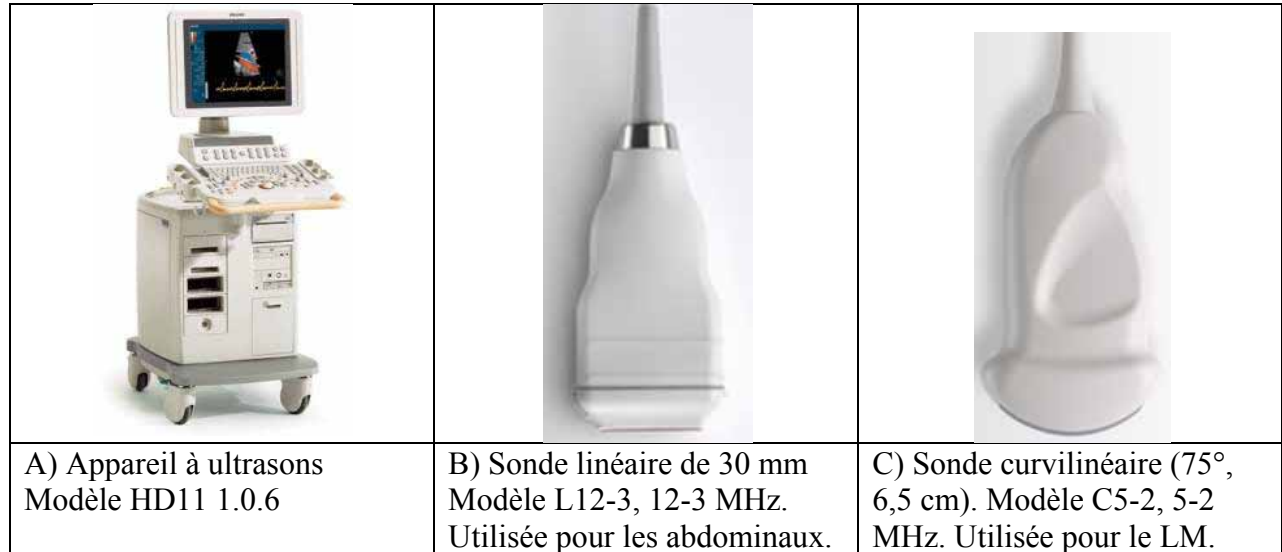


Figure 5 : Identification de l'appareil à ultrasons et de ses périphériques (sondes)

2.3.7 Traitement des images

Une image était choisie au début et à la fin de chaque vidéo. Il s'agissait de la première et de la dernière image dans presque tous les cas. Les mesures d'épaisseur du LM à deux niveaux vertébraux (L2 et L4) ainsi que des trois muscles abdominaux (TrA, OI et OE) ont été réalisées, tel que détaillé dans la section Résultats – Objectif 2 (Développement de l'interface d'analyse).

Pour le LM, la mesure d'épaisseur se faisait entre le rebord supérieur (sur l'image) de l'articulation interapophysaire (facettaire) et le rebord inférieur du fascia thoracolombaire. Pour le TrA, l'OI et l'OE, chaque mesure d'épaisseur était réalisée de façon perpendiculaire aux fascias, mais toujours en excluant ces derniers.

Deux catégories de mesures ont été considérées : les mesures primaires et les mesures dérivées. Les mesures primaires sont les mesures d'épaisseur prises au repos (mesure Repos) et lors de la contraction (mesure Contraction) alors que les mesures dérivées étaient calculées à partir de ces dernières, telles que :

- DiffRel (Différence relative) = $100 \times (\text{Contraction} - \text{Repos}) / \text{Repos}$;
- CRatio (Ratio de contraction) = $\text{Contraction} / \text{Repos}$;
- SRRatio (Ratio de symétrie au repos) = $\text{Repos muscle gauche} / \text{Repos muscle droit}$;
- SCRatio (Ratio de symétrie contraction) = $\text{Contraction muscle g.} / \text{Contraction muscle dr.}$;
- Pour les abdominaux (PTC : proportion par rapport épaisseur totale abdominaux) :

- $PTC_TrA = [TrA_Cont. / (TrA_Cont. + OI_Cont. + OE_Cont.)] \times 100$
- $PTC_OI = [OI_Cont. / (TrA_Cont. + OI_Cont. + OE_Cont.)] \times 100$
- $PTC_OE = [OE_Cont. / (TrA_Cont. + OI_Cont. + OE_Cont.)] \times 100$

Ces mesures dérivées représentent bien les différentes façons de quantifier l'atrophie ou l'activation de ces muscles, tel que rapportées dans la littérature.

2.3.8 Analyses statistiques

Toutes les analyses statistiques ont été réalisées avec l'ensemble des sujets ($n = 30$), c'est-à-dire sans tenir compte de leur statut (sujet sain ou lombalgique chronique). Cette façon de faire permet de maximiser la puissance statistique et d'obtenir des estimations (indices de fidélité) plus stables, comme l'indique une étude récente démontrant qu'un nombre minimal de sujets variant entre 25 et 30 sujets est nécessaire dans une étude de fidélité pour arriver à ces fins (Springate 2011). Les résultats correspondant à la tâche ADIM sont les seuls à faire exception, car seuls les sujets sains ($n = 15$) ont réalisé cette tâche.

La théorie générale de la fidélité (*generalizability theory*: Shavelson et Webb, 1991) a été utilisée pour estimer la fidélité des mesures ultrasonographiques. Pour être plus précis, cette théorie utilise plutôt le terme « dépendabilité » qui fait référence à la dépendance des mesures aux différentes sources d'erreurs. Cependant, le terme fidélité sera utilisé ici, car ce terme est plus commun, même s'il correspond à la théorie classique de la fidélité. Cette théorie comprend deux parties : la première est l'étude-G qui estime les composantes de variance jugées importantes pour la mesure d'intérêt (ex. : variances intersujets, interessais, interjours, interactions doubles et triples) à partir des résultats du design expérimental. Dans la présente étude, cette étude-G se traduit essentiellement en une ANOVA à deux voies avec mesures répétées sur le facteur Évaluateur (fidélité INTER-évaluateurs) ou Jour (fidélité INTRA-évaluateur) et le facteur Image. Les composantes de variance sont alors calculées à partir des valeurs de moindres carrés de l'ANOVA pour l'étude INTRA-évaluateur [Sujet (σ_s^2), Jour (σ_j^2), Image (σ_I^2), Sujet \times Jour (σ_{sj}^2), Sujet \times Image (σ_{sI}^2), Jour \times Image (σ_{jI}^2), Sujet \times Jour \times Image (σ_{sji}^2)] et INTER-évaluateurs [Sujet (σ_s^2), Évaluateur (σ_E^2), Image (σ_I^2), Sujet \times Évaluateur (σ_{sE}^2), Sujet \times Image (σ_{sI}^2), Évaluateur \times Image (σ_{EI}^2), Sujet \times Évaluateur \times Image (σ_{SEI}^2)].

La deuxième partie de la théorie générale de la fidélité est l'étude-D permettant de donner une estimation de la fidélité pour divers devis de recherche (stratégies de mesures) autres que l'étude-G. Grâce à l'étude D, la fidélité correspondant à la moyenne de 2 évaluateurs ($n_E = 2$) ou 2 jours ($n_j = 2$), et à la moyenne de 2 ou 3 essais ($n_E = 2$ ou 3), a été estimée afin d'étudier des stratégies envisageables d'un point de vue pratique. Pour l'étude de la fidélité INTER-évaluateurs, les sources de variances ont été utilisées pour calculer l'erreur standard de mesure (ESM) et l'indice de dépendabilité (ID), selon les équations suivantes :

$$ESM = \sqrt{\frac{\sigma_E^2}{n_E} + \frac{\sigma_I^2}{n_I} + \frac{\sigma_{SE}^2}{n_E} + \frac{\sigma_{SI}^2}{n_I} + \frac{\sigma_{EI}^2}{n_E n_I} + \frac{\sigma_{SEI}^2}{n_E n_I}}$$

$$\text{et}$$

$$ID = \frac{\sigma_s^2}{\sigma_s^2 + ESM}$$

où n_E et n_I sont, respectivement, le nombre d'évaluateurs et d'images considérés, lorsque des études-D sont planifiées. Pour l'étude de la fidélité INTRA-évaluateur, il faut tout simplement changer les indices « E » (Évaluateur) par « J » (Jour). Il est facilement observable à partir des deux équations illustrées ci-dessus qu'une augmentation de n_E (ou n_J) et n_I permet d'augmenter la fidélité (augmentation de ID et diminution de ESM). Dans la terminologie de la théorie générale de la fidélité, les formulations utilisées pour calculer ID et ESM correspondent aux erreurs absolues (et non aux erreurs relatives) où toutes les sources de variation sont tenues en compte. Les études G et D réalisées ici ne comprennent pas les sources de variation due au côté du corps, à la tâche, ou à l'utilisation d'un gabarit (LM) ou cube de mousse (TrA), car les stratégies de mesures étudiées (études D) ne peuvent impliquer de faire la moyenne de mesures réalisées sur différents côtés, tâches ou différentes méthodes pour contrôler la position de la sonde.

L'ID et l'ESM donnent des informations complémentaires. L'ID donne une idée de la capacité de la mesure à différencier (ordonner) les sujets et aide donc dans un processus de diagnostic. L'ESM, qui estime la variabilité de la mesure, indique la capacité à détecter un changement dans le temps (sensibilité au changement) et a donc une importance lorsque des mesures répétées sont réalisées, comme c'est le cas pour évaluer les effets de la réadaptation. À partir de l'ESM, il est possible de calculer la plus petite différence détectable (PPDD), correspondant à l'intervalle de confiance à 95 % de l'erreur de mesure ($PPDD_{95\%} = 1,96 * \sqrt{2} * ESM$). La PPDD permet de connaître la différence au-delà de laquelle un effet de traitement n'est plus attribuable à une erreur de mesure, mais à un effet thérapeutique réel. Les valeurs de PPDD ne seront pas rapportées afin de demeurer concises, mais le lecteur intéressé pourra les calculer avec les valeurs d'ESMs. Finalement, l'ESM, habituellement exprimée dans l'unité de la mesure évaluée, sera aussi exprimée en pourcentage de la moyenne de l'ensemble des mesures pour faciliter les interprétations.

Différentes analyses ANOVAs pour mesures répétées ont été réalisées afin de démontrer les effets de différentes méthodes pour améliorer la fidélité : (1) faire la moyenne de plusieurs images, (2) utiliser une tâche plutôt qu'une autre, (3) contrôler la position de la sonde (LM : gabarit; TrA : cube de mousse), (4) faire trois mesures juxtaposées pour les muscles abdominaux. En d'autres mots, les IDs et ESMs ont été comparés avec des tests paramétriques (ANOVA, test-t). Les IDs et ESMs étaient normalement distribués de sorte qu'aucune transformation (ex. : transformation de Fisher pour les coefficients de corrélations tels que les IDs) n'a été nécessaire pour rendre les distributions normales. Pour faciliter la compréhension, ces analyses seront expliquées au fur et à mesure dans la section résultats. Seules les principales analyses faisant l'objet de la présente étude, soit les analyses de fidélité INTER- et INTRA-évaluateur, sont détaillées ici.

3. RÉSULTATS ET DISCUSSION

3.1 Objectif 1 : Entraînement de notre équipe de recherche

Durant la période de quatre mois de familiarisation, plusieurs éléments ont dû être déterminés :

(1) le type de sonde à utiliser, (2) une procédure pour positionner la sonde sur les abdominaux, (3) les tâches automatisées à évaluer sur le plan de la fidélité et (4) une procédure pour quantifier les images. Ces éléments sont développés dans les prochains paragraphes.

Nous avons tout d'abord constaté que plutôt de prendre une image lors du repos et une deuxième image lors de la contraction, il valait mieux enregistrer le tout en mode vidéo, afin de permettre de sélectionner l'image la plus claire lors de la contraction. Cela s'avère être très utile dans certains cas. Aussi, une sonde curvilinéaire permettait de couvrir presque la totalité de la colonne vertébrale lombaire et donc, de ne recueillir qu'une image pour quantifier l'épaisseur du LM à plusieurs niveaux vertébraux, ce qui raccourcissait l'évaluation. Cependant, la sonde curvilinéaire produit une image déformée, ce qui devait être pris en compte lors de la quantification des images (détails dans la section suivante). Pour les abdominaux, c'est la sonde linéaire qui donnait une image plus nette.

Pour les muscles abdominaux, nous avons cherché une méthode permettant de positionner la sonde en fonction de repères anatomiques externes. Le but est d'obtenir une image permettant de réaliser des mesures à un endroit où les trois muscles abdominaux (TrA, OI, OE) ont une épaisseur relativement uniforme (bordures de chaque muscle sont parallèles). La méthode la plus citée dans la littérature recommande de placer la sonde à l'intersection de la ligne horizontale passant par le nombril et de la ligne verticale passant par le pli axillaire. Or, cette définition ne permet pas d'obtenir une image où les trois couches de muscles (TrA, OI, OE) sont parallèles chez tous les sujets, ce qui demande presque toujours un ajustement dans le positionnement de la sonde. Il apparaît qu'il y a beaucoup de variations interindividuelles portant sur la morphologie des muscles abdominaux. Nous avons investigué différentes méthodes et celle permettant d'arriver à répondre à notre critère est décrite à la Figure 3. En résumé, pour répondre à cette variabilité interindividuelle, il s'agit de trouver un repère interne (au lieu d'externe), qui est le « V » de la partie médiane du TrA et de placer ensuite la sonde par rapport à cette référence.

Une revue de la littérature permet de cibler une tâche automatisée pour chacun des muscles : (1) pour le LM, en position couchée sur le ventre, le lever d'un poids libre (5 cm au-dessus de la table d'examen) avec le bras controlatéral au muscle évalué (Kiesel *et coll.* 2007a) (**Figure 4A**), (2) pour le TrA, en position couchée sur le dos, l'ÉAJT de 5 cm au-dessus de la table d'examen (Koppenhaver *et coll.* 2009 ; Teyhen *et coll.* 2009) (**Figure 4B**). Les autres tâches évaluées ne permettaient pas, à notre avis, un niveau de standardisation suffisamment élevé pour permettre des mesures fidèles. De plus, pour le TrA, le défi est de trouver une tâche pas trop difficile où les muscles superficiels (OI et OE) seraient préférentiellement recrutés. Or, certaines études ont démontré une activation musculaire (mesures d'électromyographie de surface) très reproductible des muscles dorsaux et abdominaux superficiels (CCIC variant entre 0,84 et 0,97 et ESM entre 6 et 9 %), chez des sujets sains et lombalgiques chroniques, lors de certaines tâches sous-maximales (Dankaerts *et coll.* 2004 ; O'Sullivan *et coll.* 1998). Ces tâches ont l'avantage de pouvoir aussi se réaliser sur une table d'examen. Pour le LM, le sujet est couché sur le ventre

avec les genoux à 90° et doit lever (5 cm) la jambe controlatérale au muscle évalué (**Figure 4C**). Pour le TrA, le sujet est couché sur le dos avec les genoux à 90° (« hooklying ») et lève de 5 cm les deux jambes à la fois (**Figure 4D**). Cette dernière tâche (Élévation active jambes fléchies - ÉAJF) a l'avantage d'être plus facile que l'ÉAJT qui est parfois assez difficile à réaliser chez certains patients. Toutes ces tâches devaient être réalisées progressivement.

Nous avons discuté de la manière de quantifier les images, ce qui nous a permis de constater qu'il y a différentes façons de faire, les unes plus discutables que d'autres. Entre autres, une lecture de la littérature actuelle ne permet pas d'avoir une idée précise de la façon de faire. On ne fait jamais mention qu'il faut considérer la déformation d'une image prise avec une sonde curvilinéaire. On ne semble pas non plus trop se préoccuper de faire des mesures d'épaisseur qui sont perpendiculaires à l'axe longitudinal du muscle. Une procédure claire pour quantifier l'épaisseur des deux muscles d'intérêt est détaillée dans la prochaine section.

3.2 Objectif 2 : Développement d'une interface d'analyses

L'interface comporte les caractéristiques principales suivantes permettant :

- D'analyser les images ou de sélectionner les images à partir d'un vidéo;
- De sélectionner un groupe d'images ou de vidéos et de les analyser en groupe, ce qui rend possible une analyse à l'aveugle, car le nom des fichiers (contient information sur conditions expérimentales) n'est pas connu. Afin de permettre une analyse complète à l'aveugle, il est possible de demander de cacher la partie supérieure de l'image où certaines informations sur le sujet peuvent être indiquées;
- D'attribuer un nom de code d'évaluateur, afin de générer des fichiers de sortie permettant d'identifier l'évaluateur;
- De calibrer l'image de manière à transformer les mesures de pixels à cm;
- D'obtenir des mesures exemptes de l'effet de distorsion de l'image qui était engendrée par l'utilisation d'une sonde curvilinéaire pour le LM, c'est-à-dire en orientant les mesures en fonction du rayon de courbure de la sonde (**Figure 6**). Évidemment, la sonde était toujours positionnée de façon la plus perpendiculaire possible avec les tissus cutanés et du fascia à la hauteur de L3 avant de faire l'acquisition du vidéo;
- De réaliser trois mesures par muscle [une au centre et une à 1 cm de chaque côté (Ferreira *et coll.* 2004)], puisque les muscles de la paroi abdominale étaient relativement parallèles dans plusieurs cas (**Figure 7**);
- De réaliser des mesures bien définies pour les muscles d'intérêt (selon la littérature), mais comporte aussi des cellules vides (n = 10), laissant assez de flexibilité pour ajouter de nouvelles mesures;
- De générer des fichiers de sortie en différents formats, afin de convenir à différentes méthodes d'analyse [Fichier texte (.txt), fichier Excel (.xls) ou Fichier Matlab (.mat)].

Différents outils d'analyse plus poussés ont été implantés et sont expliqués dans un document Word (manuel d'utilisation) et des documents vidéo, tous rédigés en anglais, compte tenu de nos collaborations en cours et à venir.

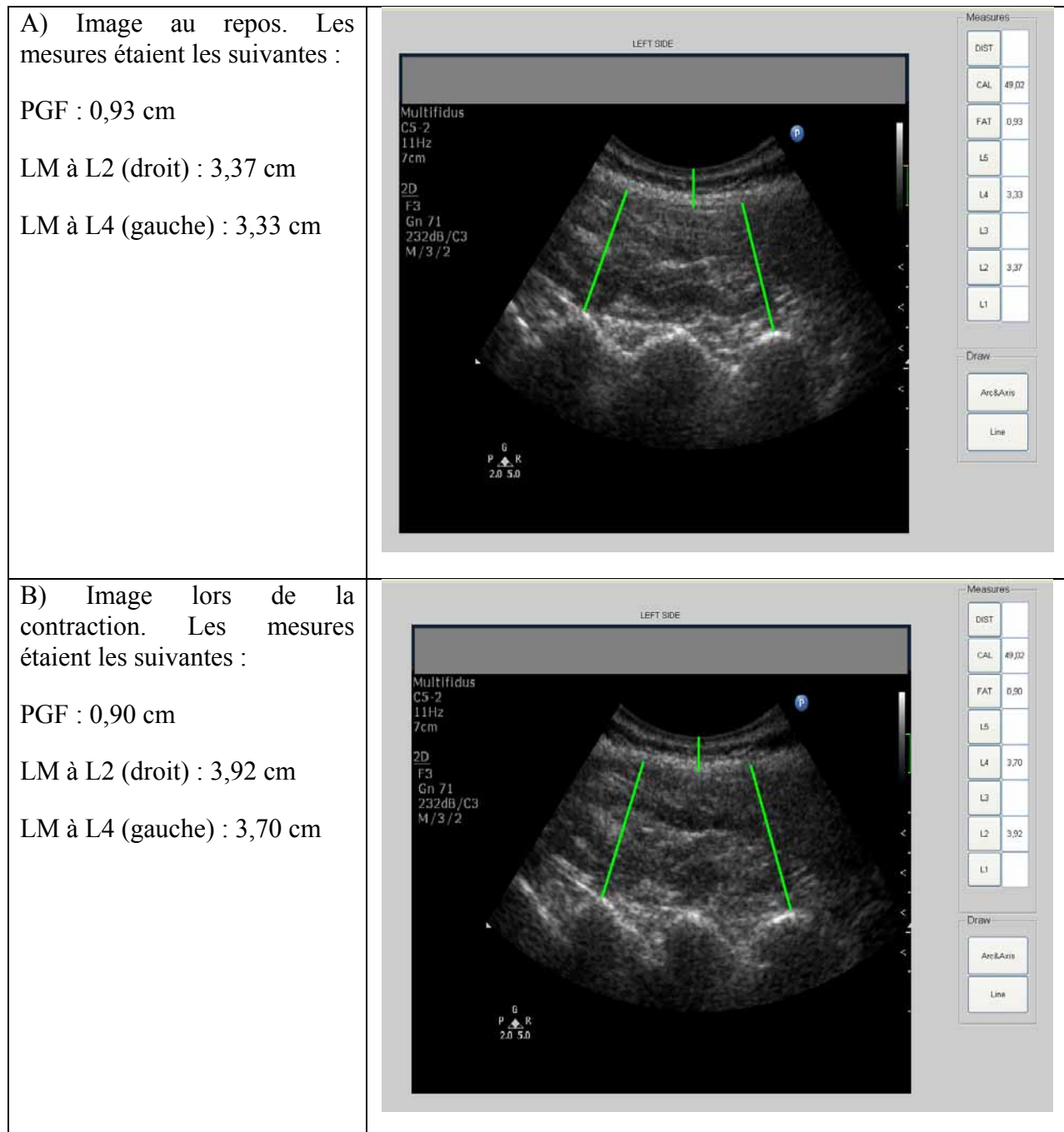


Figure 6 : Illustration des images ultrasons au repos (A) et lors de la contraction (B) du muscle multifide (LM) telles que sélectionnées et visualisées dans l’interface d’analyse Matlab.

Les lignes verticales représentent les mesures d’épaisseur obtenues pour les tissus séparant la sonde du muscle d’intérêt (PGF : Peau + Graisse + Fascia) ainsi que pour le muscle lui-même, à deux niveaux vertébraux (L2 et L4). Les outils d’analyse ainsi que le tableau pour appairer les valeurs correspondant à chaque mesure sont illustrés à droite de l’interface d’analyse. L’outil d’analyse principal permettait de déterminer le rayon de courbure de l’image (en identifiant un minimum de trois points en périphérie de la sonde) afin de dresser une ligne partant du centre du

cercle hypothétique obtenu et le bas de la figure. Cette ligne pouvait être déplacée de chaque côté afin d'aider l'évaluateur à orienter les mesures dans la bonne direction (correction de l'effet de courbure de l'image).

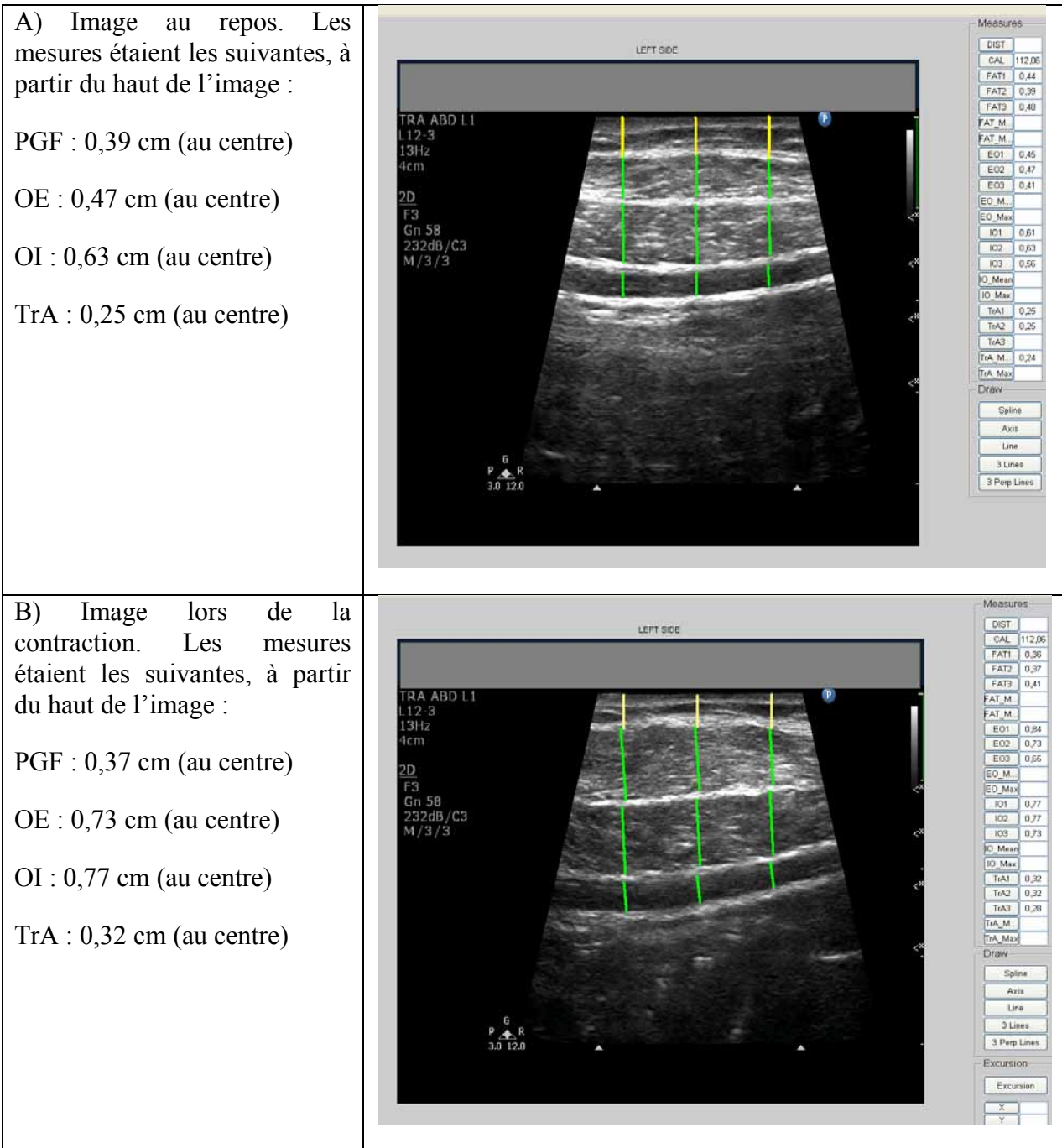


Figure 7 : Illustration des images ultrasons au repos (A) et lors de la contraction (B) du muscle transverse de l'abdomen (TrA), oblique interne (OI) et oblique externe (OE) telles que sélectionnées et visualisées dans l'interface d'analyse Matlab.

Les lignes verticales représentent les mesures d'épaisseur obtenues pour les tissus séparant la sonde du muscle d'intérêt (PGF : Peau + Graisse + Fascia) ainsi que pour les muscles eux-mêmes. Il est à noter que trois mesures juxtaposées, c.-à-d. une au centre et une à 1 Cm de chaque côté du centre, étaient prises. Les outils d'analyse ainsi que le tableau pour apparier les valeurs correspondant à chaque mesure sont illustrés à droite de l'interface d'analyse. Les outils d'analyses principaux permettaient d'orienter les mesures en fonction de l'axe longitudinal du muscle et permettaient aussi de donner des repères pour faire les mesures au centre ou à 1 cm de chaque côté du centre.

3.3 Objectif 3 : Étude de fidélité intra- et INTER-évaluateurs

Une ANOVA à 2 voies pour mesures répétées a évalué les différences entre les sessions ($n = 3$) et entre le début et la fin de chaque session en ce qui a trait à l'intensité de la douleur, telle que mesurée avec l'échelle numérique. La douleur n'était pas significativement différente ($P = 0,681$) entre la session 1 ($3,1 \pm 2,0$ cm), la session 2 ($3,4 \pm 1,8$ cm) et la session 3 ($3,2 \pm 1,4$ cm), mais démontrait une augmentation significative ($P = 0,001$), bien que modeste, entre le début ($2,8 \pm 1,7$ cm) et la fin ($3,6 \pm 1,7$ cm) des sessions de mesure. L'interaction entre ces deux facteurs était non-significative.

3.3.1 Description des données de base

Les statistiques réalisées dans cette section ont permis de comprendre s'il y avait des effets systématiques engendrés par le côté analysé et par les conditions expérimentales (tâches, contrôle de la position de la sonde) sur la contraction des différents muscles et les mesures associées, ceci dans le but d'aider à l'interprétation des résultats de fidélité. Les sources d'erreurs considérées dans l'étude de la fidélité (évaluateur, jour, image) n'ont pas été prises en compte dans ces analyses (données agrégées). De plus, seulement la mesure du centre a été utilisée pour les muscles abdominaux.

Les statistiques descriptives (moyennes et écarts-types) correspondant aux différentes mesures (Repos, Contraction, DiffRel, CRatio, SRRatio, SCRatio), muscles (LM à L4, TrA, OI, OE), côtés (gauche et droit) et aux différentes conditions expérimentales (tâches, contrôle de la position de la sonde) sont illustrées dans les **Figures 8 à 12**. On remarque l'absence d'effet du contrôle de la position de la sonde et du côté, ce qui a été confirmé avec des analyses statistiques préliminaires. Les mesures primaires du LM à L4 (Repos : $31,7 \pm 6,5$ mm; Contraction : $39,2 \pm 6,9$ mm) se comparent aux valeurs moyennes (Repos : $24,8 - 34,6$ mm; Contraction : $36,8 - 38,2$ mm) obtenues dans la littérature avec la tâche LM-bras (Kiesel *et coll.* 2007a ; Koppenhaver *et coll.* 2009 ; Wallwork *et coll.* 2007). Les mesures primaires obtenues pour le TrA lors de la tâche ÉAJT (Repos : $3,0 \pm 0,9$ mm; Contraction : $3,7 \pm 1,3$ mm) se comparent aussi aux valeurs moyennes (Repos : $3,1 - 4,4$ mm; Contraction : $3,5 - 3,7$ mm) obtenues par d'autres groupes (Koppenhaver *et coll.* 2009 ; Teyhen *et coll.* 2009). Le même constat peut être fait pour l'ADIM. Malheureusement, il n'y a pas de comparatifs pour la tâche LM-Jambes et ÉAJF (pour le TrA). La mesure DiffRel permet de constater que les tâches ont généré une contraction musculaire détectable avec la mesure par ultrasons avec des changements équivalents à 23-25 % pour LM et entre 30 et 101 % pour le TrA, selon la tâche.

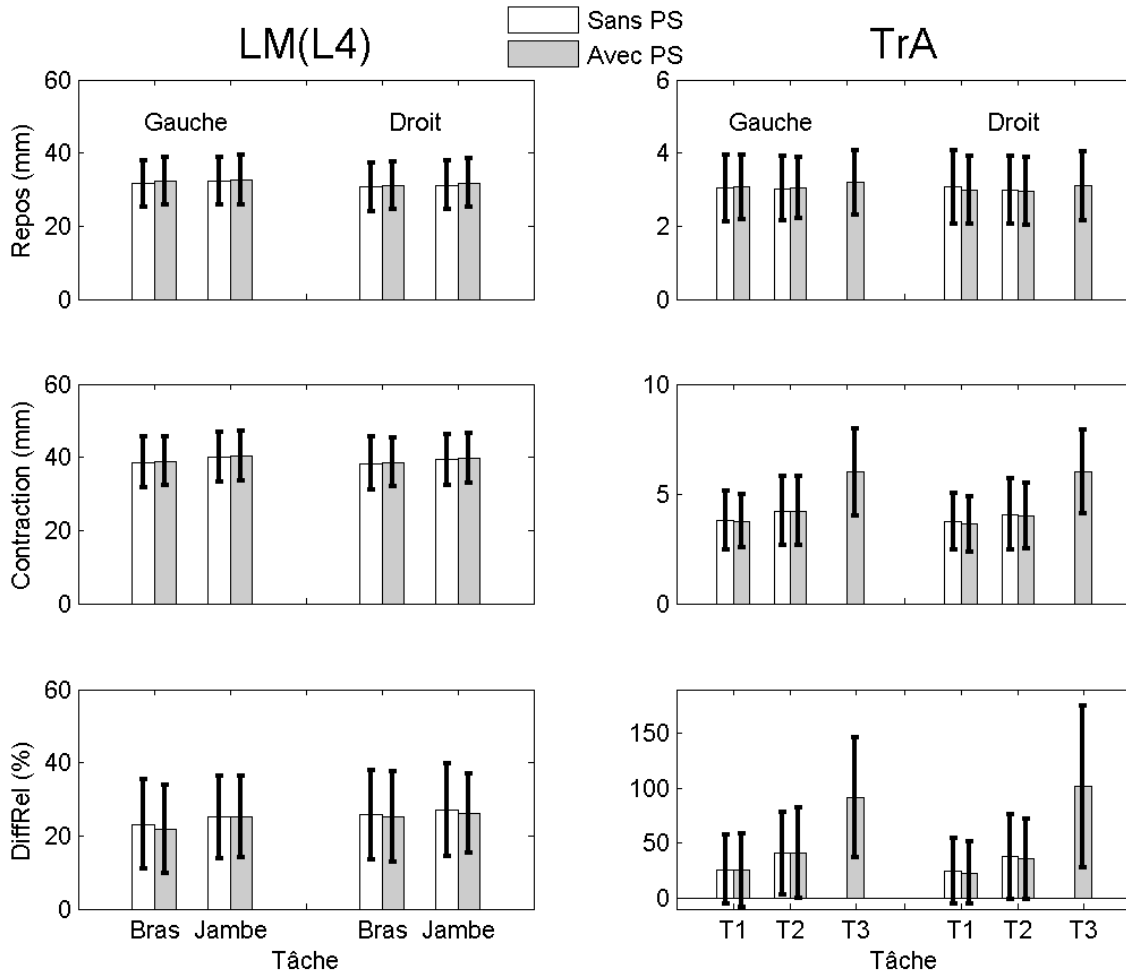


Figure 8 : Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (Repos, Contraction, DiffRel) pour les muscles LM et TrA, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM) et avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le gabarit (pour LM) ou le cube de mousse (pour TrA). Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.

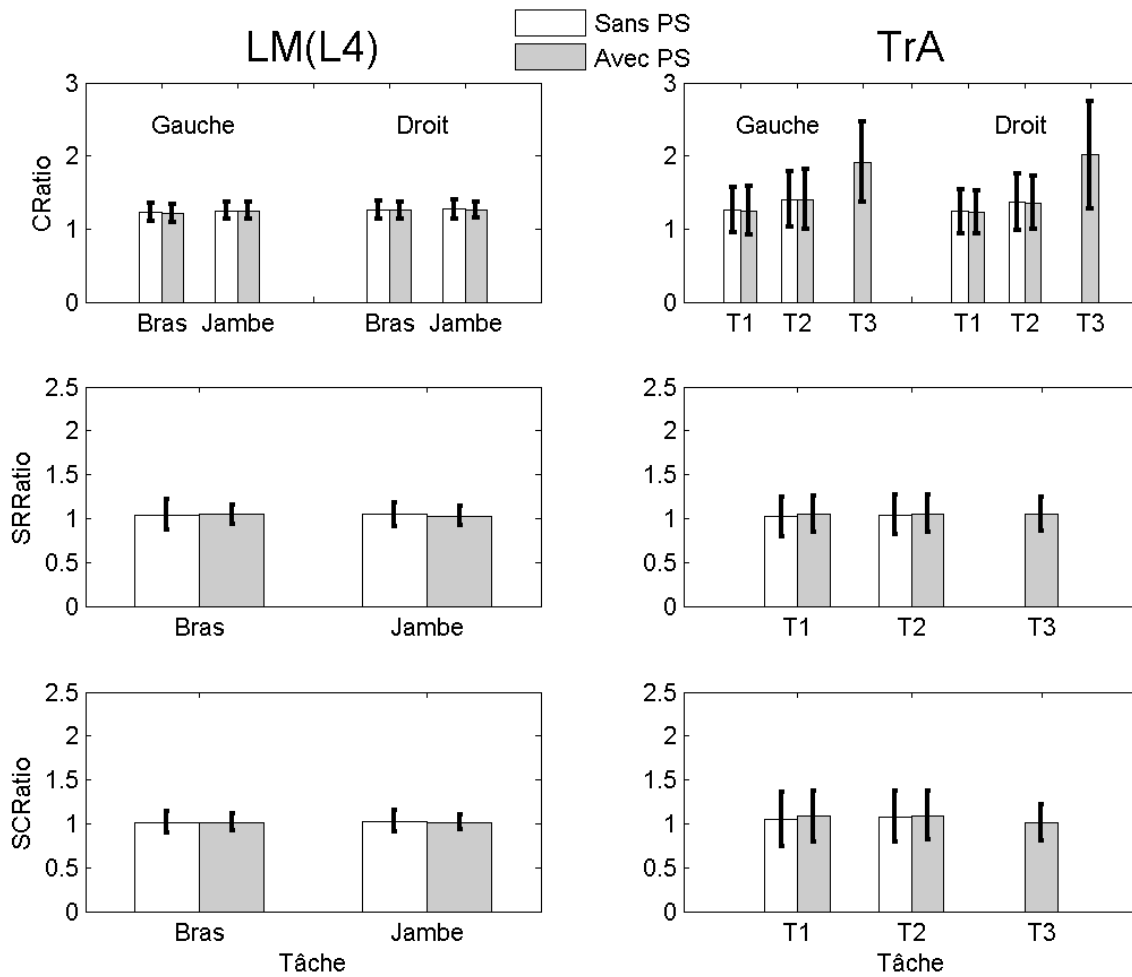


Figure 9 : Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (CRatio, SRRatio, SCRatio) pour les muscles LM et TrA, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM) et avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le gabarit (pour LM) ou le cube de mousse (pour TrA). Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.

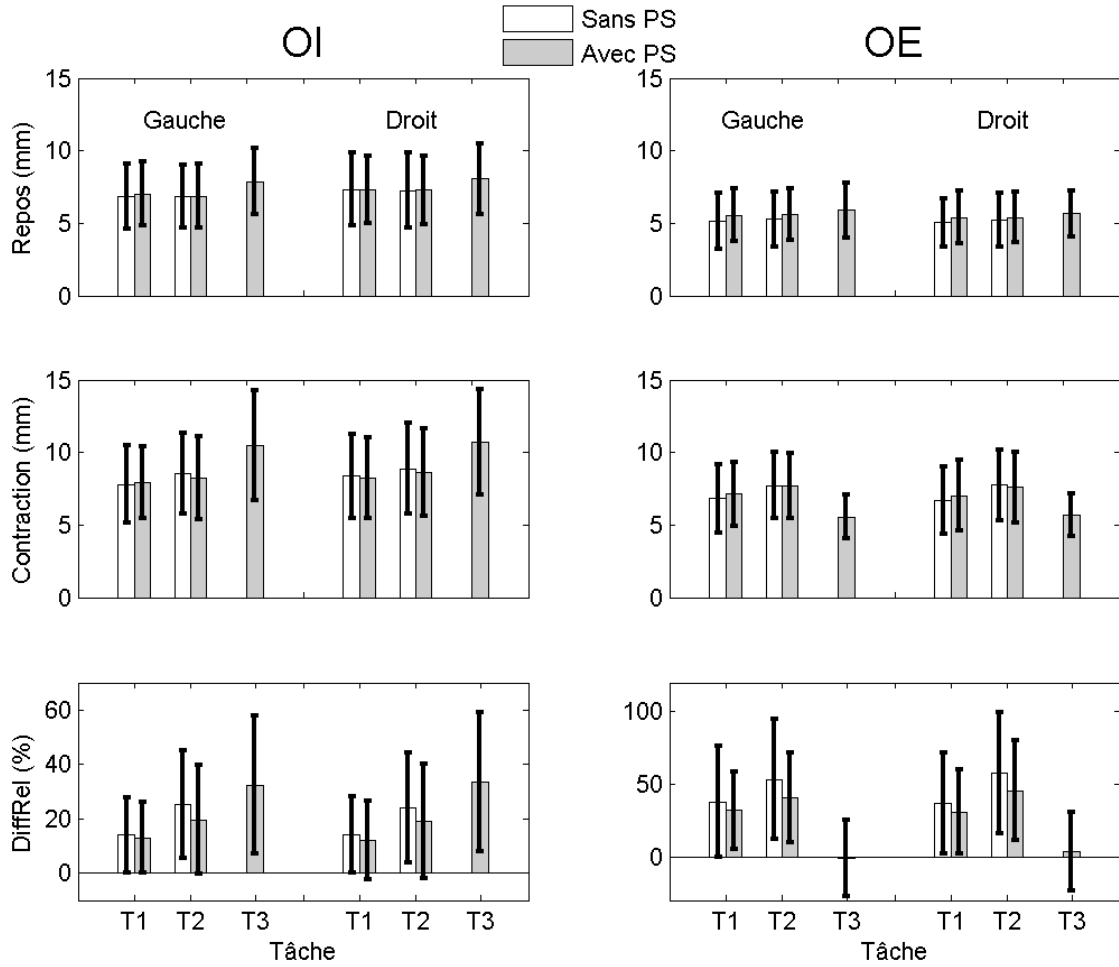


Figure 10 : Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (Repos, Contraction, DiffRel) pour les muscles OI et OE, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.

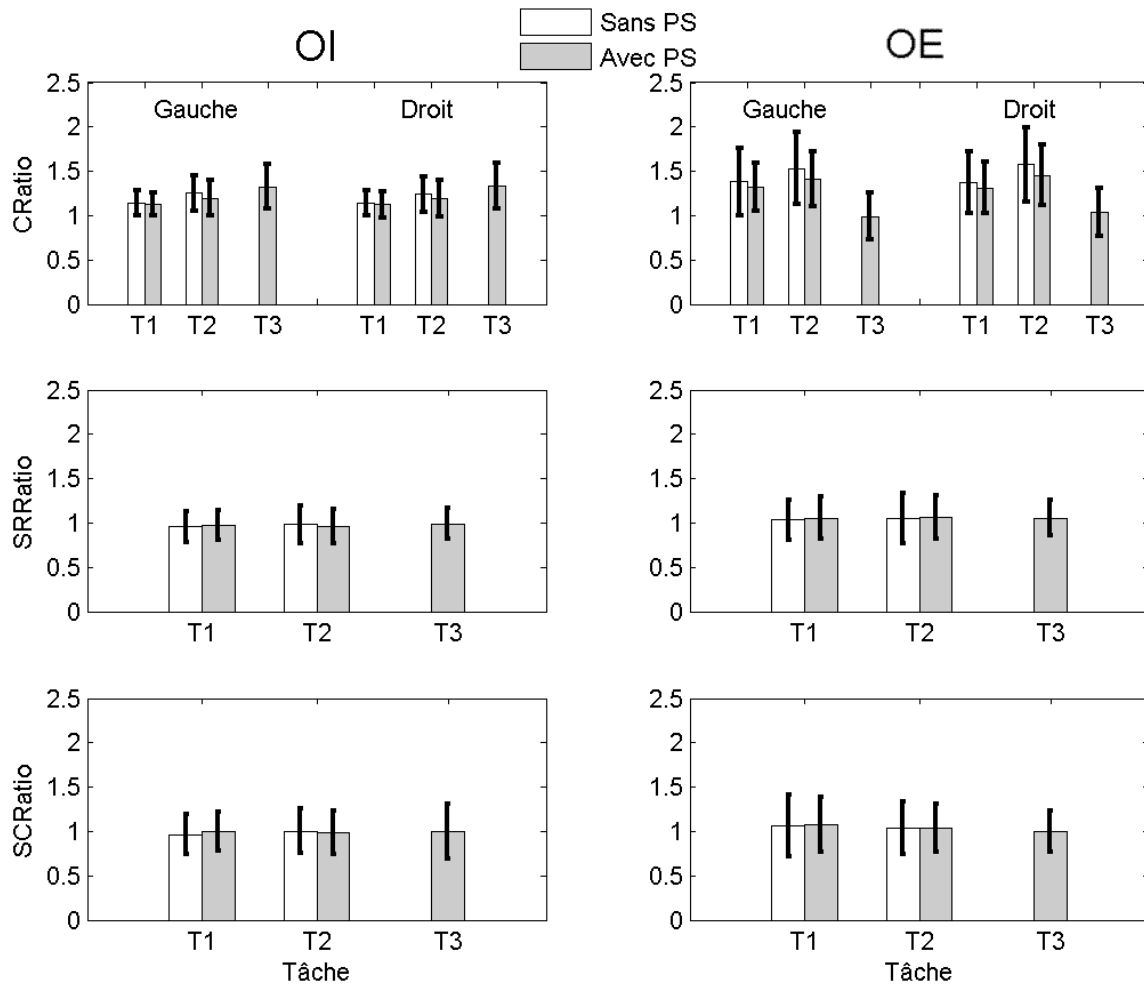


Figure 11 : Statistiques descriptives (moyennes et écarts-types) correspondant à trois mesures (CRatio, SRRatio, SCRatio) pour les muscles OI et OE, aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.

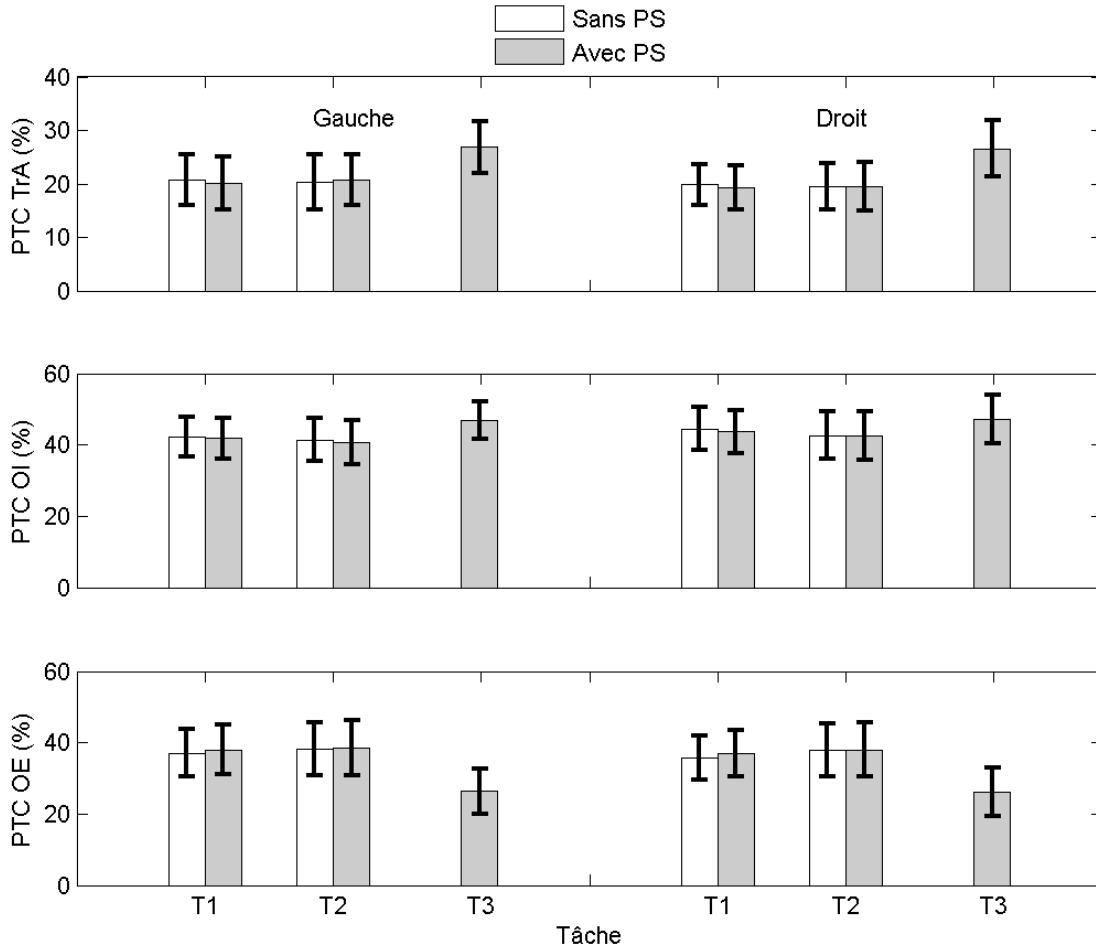


Figure 12 : Statistiques descriptives (moyennes et écarts-types) correspondant à la variable PTC (proportion par rapport épaisseur totale abdominale) pour les trois muscles abdominaux (TrA, OI et OE), aux différentes tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables ont été agrégés pour produire ces statistiques.

Pour chacun des muscles, l'épaississement du muscle lors de la contraction musculaire (Repos vs Contraction) a été comparé entre les tâches ($n = 2$ pour chaque muscle c.-à-d. sans ADIM) à l'aide d'une ANOVA à mesures répétées (2 CONTRACTIONS \times 2 TÂCHE) avec les données de l'évaluateur no 1. Les interactions CONTRACTIONS \times TÂCHE étaient significatives dans tous les cas à l'exception de LM à L2 ($P = 0,065$) où les deux tâches donnent le même résultat (**Tableau 2**). À L4, la tâche LM-jambe induisait un plus grand épaississement (8,1 mm) que LM-bras (7,2 mm), ce qui impliquait une différence de 0,9 mm (3 % de l'épaisseur au repos). TrA-ÉAJF induisait un plus grand épaississement des muscles abdominaux (TrA : 1,3 mm; OI : 1,5 mm; OE : 2,4 mm) que TrA-ÉAJT (TrA : 0,8 mm; OI : 0,9 mm; OE : 1,7 mm), des différences qui semblent être plus importantes (16, 8 et 13 % de l'épaisseur au repos, respectivement). L'effet principal CONTRACTIONS était significatif dans tous les cas, ce qui se traduisait par un épaississement de 8,2 et 7,7 mm pour LM à L2 et L4 respectivement, et un épaississement de 1,0, 1,3 et 2,0 mm pour le TrA, OI et OE respectivement.

Tableau 2 : Résultats statistiques (probabilités) correspondant aux ANOVAs (2 Contractions × 2 Tâche) appliquées aux différents muscles

Muscle	Contraction (C)	Tâche (T)	C × T
LM à L2	<0.001	0.254	0.061
LM à L4	<0.001	<0.001	0.012
TrA	<0.001	<0.001	<0.001
OI	<0.001	0.003	<0.001
OE	<0.001	0.002	0.001

Note : Les différences significatives sont identifiées en caractères gras

Une autre ANOVA à mesures répétées (3 MUSCLE × 3 TÂCHES) a été réalisée afin de déterminer quelle tâche (ÉAJT, ÉAJF, ADIM) permet de cibler plus spécifiquement le TrA relativement aux OI et OE, à l'aide de la mesure DiffRel (données des 15 sujets sains ayant aussi réalisé l'ADIM). Étant donné que l'interaction ainsi que les deux effets principaux étaient significatifs ($P < 0,05$), une ANOVA à une voie (facteur MUSCLE) a dû être menée pour chaque tâche individuellement afin de mieux comprendre la nature des effets. Tel qu'illustré à la **Figure 13**, pour les tâches ÉAJT et ÉAJF, TrA est plus sollicité (à 30 et 44 %) que l'OI (à 11 et 17 %), mais pas plus que l'OE (à 25 et 33 %). Cependant, pour l'ADIM, TrA est nettement plus sollicité (101 %) que l'OI (33 %) et l'OE (-1 %).

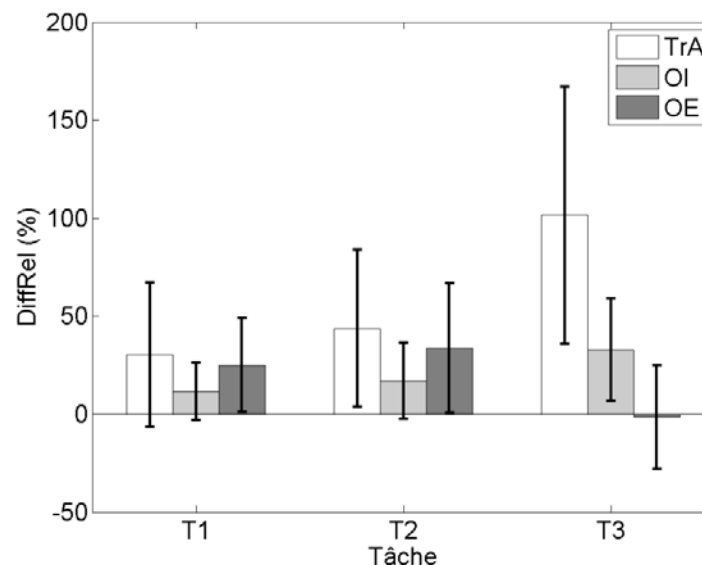


Figure 13 : Statistiques descriptives correspondant à la variable DiffRel pour les trois muscles abdominaux (TrA, OI et OE) et les trois tâches évaluées (T1 : ÉAJT; T2 : ÉAJF; T3 = ADIM), avec ou sans contrôle de la position de la sonde (avec ou sans PS) avec le cube de mousse. Les résultats associés aux autres variables (évaluateurs, jours, côtés, images) ont été agrégés pour produire ces statistiques (n = 180). Ces résultats correspondent aux 15 sujets sains seulement.

3.3.2 Étude de la fidélité

3.3.2.1 Étude G (composantes de variance)

Les résultats de l'étude G sont résumés dans les **Tableau 3** (mesures primaires : Repos et Contraction) et **Tableau 4** (mesures dérivées : DiffRel, CRatio, SRRatio, SCRatio, PTC). Ces tableaux démontrent tout d'abord que la proportion de variances attribuée aux sujets est relativement élevée pour les mesures absolues, alors qu'elle est beaucoup plus modeste pour les mesures dérivées. Les IDs de l'étude D suivent cette tendance. La principale information à retirer de ces résultats est d'identifier où sont les sources d'erreurs, afin de planifier des stratégies de mesures efficaces pour améliorer la fidélité. Bien sûr, l'interaction triple (σ_{SEI}^2 ou σ_{SJI}^2) représente une partie non négligeable de la variance. Il faut savoir que cette composante de variance contient aussi l'erreur aléatoire ainsi que les sources d'erreurs inconnues (autres que Jour, Image ou Évaluateur) n'ayant pas été prises en compte dans l'étude G. Il faudrait identifier ces dernières afin d'agir sur celles-ci. D'ailleurs, l'utilisation d'un gabarit (LM) et d'un cube de mousse (TrA) représente des solutions possibles pour diminuer l'erreur aléatoire (voir résultats plus loin). Cependant, ce sont les interactions doubles qui indiquent où les efforts de mesure devraient d'abord être concentrés et ici, comme c'est généralement le cas avec ce type de mesures, ce sont les variances σ_{SI}^2 (évaluation INTRA-évaluateur) ou σ_{EI}^2 (évaluation INTER-évaluateurs) qui se sont démarquées. Ces sources de variance indiquent que certains sujets obtiennent des mesures différentes alors que d'autres obtiennent des mesures plus semblables entre évaluateurs (fidélité INTER-évaluateurs) ou entre les jours (fidélité INTRA-évaluateur). Il est donc possible de prédire que faire la moyenne de plus d'un évaluateur (fidélité INTER-évaluateurs) ou de plus d'un jour (fidélité INTRA-évaluateur) sera une stratégie plus efficace pour améliorer la fidélité que de faire la moyenne de plus d'une image (car la variance σ_{SI}^2 est basse), ce qui sera vérifié avec l'étude D (prochaine section).

Les probabilités (valeurs P) associées à chaque étude G (ANOVAs pour mesures répétées) ont aussi permis de vérifier s'il y avait présence d'erreurs systématiques entre les évaluateurs, les jours et les images. Si nous nous attardons aux variables primaires, le nombre de différences significatives était relativement faible pour la fidélité INTRA-évaluateur autant pour LM (facteur Jour : 3/32 cas; facteur Image : 5/32 cas) que pour TrA (facteur Jour : 1/20 cas; facteur Image : 4/20 cas). Par contre, ce n'était pas le cas pour la fidélité INTER-évaluateurs, notamment pour le facteur Évaluateur, ceci autant pour LM (facteur Évaluateur : 18/32 cas; facteur Image : 7/32 cas) que pour TrA (facteur Évaluateur : 7/20 cas; facteur Image : 4/20 cas). Pour le TrA, la proportion de variance Évaluateur (σ_E^2) correspondant à ces différences significatives variait entre 3 et 7 %, ce qui implique des différences moyennes de l'ordre de 0,2 à 0,7 mm (**Figure 14**), ce qui comparable aux différences (entre 0,2 et 0,4 mm) rapportées ailleurs (Koppenhaver *et coll.* 2009). Pour LM, les valeurs de variance σ_E^2 correspondant aux mesures à L2 variaient entre 2 et 6 % alors que pour les mesures à L4, elles variaient entre 4 et 18 %. Les résultats significatifs correspondant au LM à L4 impliquent des différences moyennes de l'ordre de 2,8 à 4,1 mm (**Figure 15**), ce qui est plus élevé que les différences (entre 2,3 et 2,4 mm) rapportées ailleurs (Koppenhaver *et coll.* 2009). Ces biais ne peuvent s'expliquer par l'apprentissage des sujets ou par le temps de la journée (a.m. ou p.m.) où l'évaluation a été réalisée, car l'ordre des évaluateurs a été contrebalancé à travers les sujets. La seule autre étude ayant fait une évaluation des biais

entre évaluateurs autant pour TrA que pour LM (Koppenhaver *et coll.* 2009) avait aussi obtenu une excellente fidélité INTER-évaluateurs pour quantifier la même image (CCIC > 0.96). Ces auteurs concluent donc que l'erreur entre évaluateurs s'introduit principalement lors de l'acquisition des images. Pour le TrA, il est possible qu'un biais se soit introduit dans l'identification du V du TrA, à partir duquel la position latérale de la sonde est déterminée (3 cm latéralement). En effet, les différences significatives obtenues entre les trois mesures parallèles suggèrent qu'un tel biais de positionnement de la sonde pourrait expliquer ces différences INTER-évaluateurs. Pour le LM, on pourrait penser que les évaluateurs pourraient se tromper d'un niveau vertébral, de façon systématique, mais le fait que l'utilisation du gabarit n'ait pas amélioré la situation nous permet de rejeter cette possibilité. Comme mentionné en introduction, l'utilisation d'un cube de mousse n'est pas possible pour contrôler la pression et l'inclinaison appliquées sur la sonde, non seulement à cause des formes irrégulières près de la colonne vertébrale, mais aussi parce qu'il faut faire varier l'inclinaison de la sonde pour faire apparaître les apophyses transverses nécessaires à la quantification du LM (Stokes *et coll.* 2005). D'ailleurs, même si une image nette était obtenue au début de chaque vidéo (mesure au repos), l'augmentation de l'épaisseur du muscle lors de la contraction induisait tellement de variations dans les images qu'il fallait changer l'orientation initiale de la sonde une fois que la contraction était complétée de manière à mieux voir ou même parfois retrouver les facettes qui avaient disparu. Changer l'angle d'inclinaison a le potentiel de changer la position visible des facettes sur l'image, ce qui peut se traduire en erreurs de mesures, comme déjà évalué pour le TrA (Whittaker *et coll.* 2009). Cette étude n'a pas relevé de changement significatif de l'épaisseur du TrA avec des changements d'inclinaison inférieure à 9°, soit la plus grande variation d'inclinaison évaluée. Cependant, ces résultats peuvent ne pas être applicables au LM et les changements d'inclinaison induits dans la présente étude peuvent avoir été supérieurs à 9°. Il est possible que les évaluateurs aient utilisé des angles d'inclinaison systématiquement différents, ce qui expliquerait les biais entre évaluateurs. Si c'est le cas, cela pourrait aussi induire des différences inter-évaluateur plus importantes à L4 qu'à L2 si on considère que les facettes sont situées plus en profondeur à L4, comme le suggèrent les valeurs d'épaisseurs plus élevées à L4 (≈ 33 mm, Tableau A-4) qu'à L2 (≈ 28 mm, Tableau A-7). Contrairement à la présente étude, une autre étude ne démontre pas de différence en termes de fidélité, entre les niveaux L2/L3 et L4/L5 (Wallwork *et coll.* 2007). Il est dommage que les mesures à L4 aient été moins fidèles dans la présente étude, car l'atrophie du LM se fait surtout aux niveaux L4 et L5, comparativement aux niveaux lombaires supérieurs (Hyun *et coll.* 2007 ; Wallwork *et coll.* 2008). Une récente étude a aussi démontré que suite à une période de repos au lit (60 jours), l'atrophie se fait aussi plus rapidement à ces niveaux et permet même de prédire l'occurrence d'un nouvel épisode de lombalgie (Belavy *et coll.* 2010). Il est clair que des procédures innovatrices doivent être proposées et évaluées pour réduire les potentiellement importantes sources de variation induites par un mauvais contrôle de l'inclinaison de la sonde, en plus de permettre un bon contrôle de la pression appliquée sur la peau.

Tableau 3 : Pourcentages attribués à chaque composante de variance, selon l'étude G : résultats pour les mesures primaires (au repos et lors de la contraction)

Muscle	Fidélité	Statistiques*	Composantes de variance (% variance totale)						
			Sujet (S)	Jour (J)	Image (I)	S×J	S×I	J×I	S×J×I
LM	INTRA- évaluateur	(n = 32 études G)							
		Moyenne	88	0	0	5	1	0	5
		5 ^e percentile	82	0	0	2	0	0	2
		95 ^e percentile	93	1	0	9	5	0	8
LM	INTER- évaluateurs	(n = 32 études G)	Sujet (S)	Éval. (E)	Image (I)	S×E	S×I	E×I	S×E×I
		Moyenne	69	6	0	17	1	0	7
		5 ^e percentile	52	0	0	6	0	0	4
		95 ^e percentile	83	15	0	36	2	1	13
TrA	INTRA- évaluateur	(n = 20 études G)	Sujet (S)	Jour (J)	Image (I)	S×J	S×I	J×I	S×J×I
		Moyenne	80	0	0	11	0	1	7
		5 ^e percentile	73	0	0	5	0	0	3
		95 ^e percentile	88	1	0	19	2	3	14
TrA	INTER- évaluateurs	(n = 20 études G)	Sujet (S)	Éval. (E)	Image (I)	S×E	S×I	E×I	S×E×I
		Moyenne	73	2	0	13	1	1	11
		5 ^e percentile	65	0	0	4	0	0	6
		95 ^e percentile	79	7	0	22	2	4	16

* Pour LM, ces statistiques descriptives ont été dérivées de 32 études G [2 mesures (Repos, Contraction) × 2 Côtés (gauche et droit) × 2 Tâches (soulèvement bras et jambe) × 2 Gabarits (sans et avec) × 2 Niveaux (L2 et L4)]. Pour TrA, elles ont été dérivées de 20 études G [2 mesures (Repos, Contraction) × 2 Côtés (gauche et droit) × 2 Tâches (ÉAJT et ÉAJF) × 2 Cubes (sans et avec) auxquelles sont ajoutées 2 mesures × 2 côtés × 1 autre tâche (ADIM)].

Éval : évaluateur

Tableau 4 : Pourcentages attribués à chaque composante de variance, selon l'étude G : résultats pour les mesures dérivées (DiffRel, CRatio, SRRatio, SCRatio; PTC pour les muscles abdominaux)

Muscle	Fidélité	Statistiques*	Composantes de variance (% variance totale)						
			Sujet (S)	Jour (J)	Image (I)	S×J	S×I	J×I	S×J×I
LM	INTRA- évaluateur	(n = 64 études G)	Sujet (S)	Jour (J)	Image (I)	S×J	S×I	J×I	S×J×I
		Moyenne	41	1	1	18	5	1	33
		5 ^e percentile	19	0	0	0	0	0	20
		95 ^e percentile	58	3	4	42	16	3	52
LM	INTER- évaluateurs	(n = 64 études G)	Sujet (S)	Éval. (E)	Image (I)	S×E	S×I	E×I	S×E×I
		Moyenne	24	2	0	37	2	2	32
		5 ^e percentile	7	0	0	18	0	0	15
		95 ^e percentile	40	9	3	62	10	12	51
TrA	INTRA- évaluateur	(n = 50 études G)	Sujet (S)	Jour (J)	Image (I)	S×J	S×I	J×I	S×J×I
		Moyenne	47	0	0	23	2	2	25
		5 ^e percentile	5	0	0	6	0	0	10
		95 ^e percentile	71	2	1	44	8	6	51
TrA	INTER- évaluateurs	(n = 50 études G)	Sujet (S)	Éval. (E)	Image (I)	S×E	S×I	E×I	S×E×I
		Moyenne	38	1	0	23	2	2	34
		5 ^e percentile	0	0	0	0	0	0	13
		95 ^e percentile	63	7	1	50	8	11	49

* Pour LM, ces statistiques descriptives ont été dérivées de 64 études G [4 mesures (DiffRel, CRatio, SRRatio, SCRatio) × 2 Côtés (gauche et droit) × 2 Tâches (soulèvement bras et jambe) × 2 Gabarits (sans et avec) × 2 Niveaux (L2 et L4)]. Pour TrA, elles ont été dérivées de 50 études G [4 mesures (DiffRel, CRatio, SRRatio, SCRatio, PTC) × 2 Côtés (gauche et droit) × 2 Tâches (ÉAJT et ÉAJF) × 2 Cubes (sans et avec) auxquelles sont ajoutées ces mêmes 4 mesures × 2 côtés × 1 autre tâche (ADIM)].

Éval : évaluateur

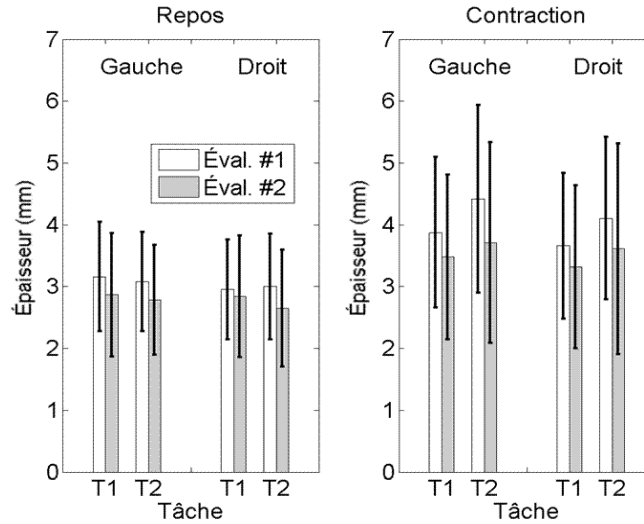


Figure 14 : Illustration des conditions expérimentales qui ont généré des différences significatives (ANOVAs pour mesures répétées; $P < 0,05$) entre les évaluateurs pour le muscle TrA (conditions expérimentales avec le cube de mousse). T1 : ÉAJT; T2 : ÉAJF

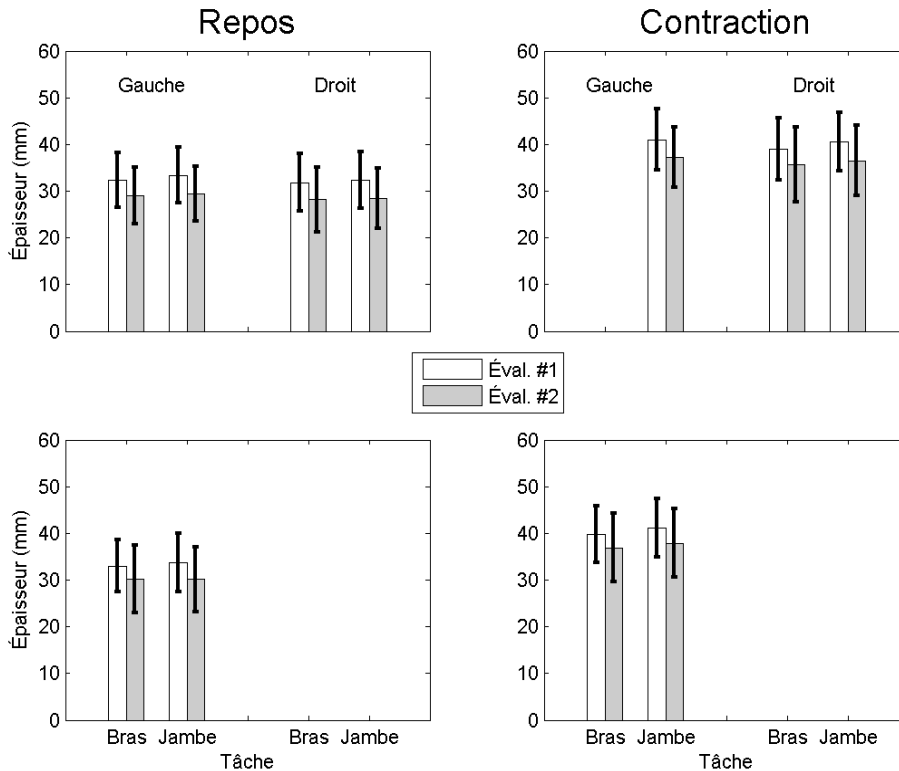


Figure 15 : Illustration des conditions expérimentales qui ont généré des différences significatives (ANOVAs pour mesures répétées; $P < 0,05$) entre les évaluateurs pour le muscle LM à L4. Les figures du haut et du bas correspondent aux conditions sans (haut) et avec (bas) utilisation d'un gabarit (acétate).

3.3.2.2 Étude D (stratégies de mesure)

L'étude G suggère que la stratégie de mesure la plus efficace serait de faire la moyenne de plus d'un évaluateur ou de plus d'un jour d'évaluation pour rehausser la fidélité INTER-évaluateurs et intra-évaluateur, respectivement. Les résultats de l'étude D supportent ces prédictions (**Tableau 5**). Cependant, cette stratégie est coûteuse en termes de ressources et exigeante pour les patients, ce qui fait qu'elle risque fort bien de ne pas être utilisée, autant en clinique qu'en recherche. C'est d'ailleurs pourquoi la stratégie la plus utilisée est de faire la moyenne de plus d'un essai (images) réalisés lors de la même évaluation. Les résultats du **Tableau 5** démontrent que l'amélioration relative de la fidélité est non négligeable lorsque deux images sont considérées (augmentation des IDs de 3 à 6 % et diminution de l'ESM de 7 à 14 %). Bien que l'analyse d'une troisième image ne semble pas beaucoup améliorer la situation (augmentation de l'ID de 1 à 3 % et diminution de l'ESM de 3 à 6 %), cette amélioration prend tout son sens lorsque la PPDD est considérée, car cette PPDD ($PPDD = 1.96 * \sqrt{2} * ESM$) se voit réduite de près de 11 % pour une diminution de 4 % de l'ESM. Conséquemment, seuls les résultats détaillés correspondant aux deux stratégies de mesure envisageables (1 ou 3 images) sont présentés à l'Annexe A (**Tableau A-1 à A-6**). Afin de réduire le nombre de tableaux, seuls les résultats des muscles LM (à L4) et TrA du côté gauche (résultats similaires du côté droit), ceci pour les neuf mesures expliquées dans la méthodologie, y sont présentés.

L'analyse des tableaux de l'annexe A permet d'observer que la fidélité INTRA-évaluateur est plus élevée que la fidélité INTER-évaluateurs, comme attendu et observé antérieurement pour ce type de mesures (Koppenhaver *et coll.* 2009 ; Watanabe *et coll.* 2004). Cependant, il faut considérer que les estimations de la fidélité INTER-évaluateurs risquent d'être plus conservatrices que dans les études antérieures, car les deux évaluateurs n'ont pas pris les mesures dans la même séance. Nous croyons que les résultats de la présente étude sont plus généralisables à ce qui serait fait en pratique (en recherche ou en clinique), car les mesures réalisées par plus d'un évaluateur sur un même sujet ne risquent pas d'être prises à quelques minutes d'intervalles.

On peut aussi y remarquer dans les tableaux de l'annexe A que les mesures primaires (Repos, Contraction) mènent à une meilleure fidélité en termes de valeurs d'IDs, que les mesures dérivées. Ces résultats sont en accord avec les études antérieures (Koppenhaver *et coll.* 2009 ; Mannion *et coll.* 2008 ; Rankin *et coll.* 2006) et ne semblent pas s'expliquer par le fait que les mesures dérivées combinent les erreurs de plus d'une mesure primaire, car les ESM sont comparables. Les mesures dérivées représentent une forme de normalisation, ce qui a pour effet de réduire la variance intersujet introduite par les différences individuelles en termes de gabarit, notamment en ce qui a trait aux différences entre les sexes (Marras *et coll.* 2001 ; Rankin *et coll.* 2006). D'ailleurs, des corrélations significatives ($P < 0,001$) ont été obtenues entre la masse et la mesure Repos pour LM ($r = 0,63$ lorsque les 2 niveaux sont regroupés), TrA ($r = 0,46$), OI ($r = 0,63$) et OE ($r = 0,49$). Les corrélations avec la variable Contraction sont moins élevées, mais toutes significatives. L'ID, comme c'est aussi le cas pour le CCIC, dépend beaucoup de la variabilité intersujet alors lorsque cette variabilité est réduite, l'ID ou le CCIC décroît aussi.

Tableau 5 : Effet de différentes stratégies de mesures (résultats de l'étude D) sur l'amélioration de la fidélité : Pourcentage d'augmentation (ID) ou de diminution (ESM) des indices de fidélité par rapport à la stratégie minimale (référence : 1 image)*

Muscle	Mesure	Fidélité	Référence		Stratégies de mesures : nombre évaluateurs/jours – nombre images †									
			(1X-1I)		1X-2I		1X-3I		2X-1I		2X-2I		2X-3I	
			ID	ESM (mm)	↑ ID	↓ ESM	↑ ID	↓ ESM	↑ ID	↓ ESM	↑ ID	↓ ESM	↑ ID	↓ ESM
LM	Repos	INTRA	0,86	2,2	4	15	5	21	6	25	9	38	10	42
	Cont.	INTRA	0,90	2,1	3	14	3	19	5	27	6	38	7	42
	Repos	INTER	0,66	3,7	5	7	6	10	20	29	23	34	25	36
	Cont.	INTER	0,72	3,8	4	7	5	10	16	29	18	34	19	36
TrA	Repos	INTRA	0,82	0,4	4	11	5	15	10	28	12	37	13	40
	Cont.	INTRA	0,79	0,7	5	12	7	16	11	28	15	37	16	41
	Repos	INTER	0,72	0,5	6	12	9	16	15	29	20	37	21	40
	Cont.	INTER	0,74	0,8	6	12	8	16	14	28	19	37	20	41

* Résultats moyens à travers 16 études G pour le LM (2 côtés × 2 tâches × 2 gabarits × 2 niveaux vertébraux) ou 10 études G pour le TrA [2 côtés × 2 tâches × 2 cubes + 2 côtés × 1 tâche (ADIM) × 1 cube]

† Exemple : 2X-3I veut dire 2 évaluateurs – 3 images dans le cas de la fidélité INTER-évaluateurs et 2 jours – 3 images dans le cas de la fidélité INTRA-évaluateur

LM : multifide lombaire; TrA : transverse de l'abdomen; Cont. : Contraction; INTRA- et INTER : fidélité intra- et INTER-évaluateurs; ID : Indice de dépendabilité; ESM : erreur standard de mesure

Les comparaisons avec la littérature sont très difficiles à réaliser étant donné que la présente étude est la seule à avoir utilisé la théorie générale de la fidélité et le calcul de l'ID qui lui est associé. Les études antérieures ont toutes considéré des CCIC de type 2 ou 3 (Shrout and Fleiss 1979). Dans l'ensemble, les résultats de la présente étude semblent se situer légèrement sous les valeurs rapportées dans la littérature. Cependant, il y a beaucoup de variations dans les devis de recherche et les statistiques utilisées. Il est possible de cibler des études qui se comparent à la méthodologie utilisée ici et de comparer les résultats (**Tableau 6**), mais tout en ayant en tête que des CCIC ont été utilisés dans cette littérature. De plus, il faut considérer que toutes les études INTER-évaluateurs comprenaient des mesures prises à l'intérieur d'une même séance, ce qui risque de favoriser la fidélité. En effet, le fait d'avoir réalisé les mesures sur deux séances de mesures différentes, même si ces séances étaient dans la même journée, a fort probablement introduit une source d'erreur supplémentaire. Cependant, à notre avis ces conditions de mesures sont plus généralisables à la pratique clinique, où les évaluateurs réaliseraient des mesures sur des sessions différentes, que ce qui a été considéré dans les études antérieures.

3.3.2.2.1 Effet de la tâche et du contrôle du positionnement de la sonde sur la fidélité

Autant pour le LM (à L2 et L4) que pour les abdominaux (TrA, OI, OE), la tâche n'a eu aucun effet sur la fidélité (IDs et ESMs comparables) INTRA- et INTER-évaluateurs (**Tableau 7**).

Pour le LM (à L2 et L4), l'utilisation du gabarit pour repositionner la sonde a mené à des IDs significativement ($P < 0,05$) plus élevées ($0,89 \pm 0,03$) et des ESMs significativement plus petites ($6,1 \pm 1,4 \%$) comparativement au fait de ne pas utiliser cette méthode (IDs : $0,87 \pm 0,04$; ESMs : $6,8 \pm 1,9 \%$) (**Tableau 7**). Étant donné que la constitution d'un gabarit requiert un certain temps (environ 10 min.), nous proposons que cette faible amélioration de la fidélité ne justifie pas cette pratique.

Pour les abdominaux (TrA, OI, OE), l'utilisation du cube de mousse a eu pour effet de diminuer significativement ($P < 0,05$), mais de façon marginale (1,2 %), les ESMs (pas pour les IDs) associées à la fidélité INTER-évaluateurs (**Tableau 7**). Aucun effet n'a été obtenu pour la fidélité INTRA-évaluateur. Encore ici, cette pratique semble ne pas être aussi efficace que souhaité. Par contre, la méthode est très simple, comparativement à l'utilisation d'un gabarit. Nous recommandons donc l'utilisation du cube de mousse.

Les résultats de fidélité correspondant au LM à L2 (côté gauche) sont résumés dans un seul tableau (Annexe A – **Tableau A-7**), pour les principales mesures (Repos, Contraction, CRatio) et les deux stratégies de mesure envisageables (1 ou 3 images). Ces résultats sont généralement comparables au LM à L4 (résultats répartis dans les Tableaux A-1, A-2, A-4 et A-5) et même supérieurs en ce qui a trait à la fidélité INTER-évaluateurs des variables Repos et Contraction. Ces derniers résultats s'expliquent par les différences systématiques observées entre les évaluateurs, qui étaient plus importantes à L4 (σ_E^2 entre 7 et 17 %) qu'à L2 (σ_E^2 entre 2 et 6 %), comme l'indiquaient les résultats de l'étude G.

Tableau 6 : Comparaison des indices de fidélité entre la présente étude et les étendues de valeurs rapportées dans la littérature sur le sujet*

Muscle	Étude	Sujets	Type de variable	INTRA-évaluateur (jours différents)		INTER-évaluateurs (images différentes)		
				CCIC ou ID	ESM (mm) †	CCIC ou ID	ESM (mm) †	
LM	Présente étude (Koppenhaver <i>et coll.</i> 2009)	15 cont. +	Repos - Contraction	0,90-0,96	1,1-1,9	0,62-0,83	2,9-4,2	
		15 lomb.	DiffRel	0,54-0,71	4,8-5,8	0,09-0,57	7,2-10,3 %	
		30 lomb.	Repos - Contraction	0,97-0,98	0,9-1,0	0,88-0,92	1,8-2,1	
			DiffRel	0,79	4,0	0,73	5,5 %	
	(Wallwork <i>et coll.</i> 2007)	10 cont.	Repos	/	/	0,97	5	
TrA	Présente étude (Koppenhaver <i>et coll.</i> 2009)		Repos - Contraction	0,77-0,92	0,2-0,7	0,75-0,83	0,4-0,8	
			DiffRel	0,62-0,84	14,8-20,5	0,43-0,69	16,5-23,9 %	
		30 lomb.	Repos - Contraction	0,93-0,94	0,2-0,4	0,80-0,91	0,3-0,6	
			DiffRel	0,89	12,3	0,78	15,6 %	
		(Mannion <i>et coll.</i> 2008)	14 cont. + 14 lomb.	Repos	0,63-0,89	0,27-0,46		
		(Norasteh <i>et coll.</i> 2007)	15 lomb.	Repos	/	/	0,80	0,45
	(Teyhen <i>et coll.</i> 2009)	15 cont. + 15 lomb.	Repos - Contraction	/	/	0,94-0,96	0,25-0,27	
			DiffRel	/	/	0,88	7,1 %	

Cont : sujets contrôles (sains) ; Lomb : sujets lombalgiques

* Les comparaisons se concentrent sur les éléments les plus rapportés dans la littérature, soit les muscles LM à L4 (Tableau A-4) et TrA, les variables Repos-Contraction et DiffRel, et pour les tâches automatisées seulement (pas l'ADIM). Les valeurs correspondant à la moyenne de trois images (présente étude) sont comparées à la moyenne de 2 ou 3 images (autres études). Les valeurs de la présente étude représentent l'étendue des valeurs à travers les 2 tâches automatisées, peu importe le niveau de contrôle de la position de la sonde. Les études choisies sont les plus récentes revues de la littérature rapportant des étendues de valeurs, auxquelles sont ajoutées des études ayant des données plus complètes et spécifiques. La fidélité INTRA-évaluateur devait être réalisée sur des jours différents (CCIC interjours). La fidélité INTER-évaluateurs devait porter sur des images différentes (recueillies par des évaluateurs différents).

† Valeurs en mm à l'exception de DiffRel.

Tableau 7 : Résultats statistiques (valeurs *P*) correspondant aux Test-t (pour données paires) pour comparer les IDs et les ESMs correspondant à différentes tâches et au contrôle du positionnement de la sonde.

Type de fidélité	LM à L2 et L4*		TrA-OI-OE†	
	ID	ESM	ID	ESM
Effet de la tâche				
INTRA-évaluateur	0,094	0,217	0,879	0,648
INTER-évaluateurs	0,154	0,143	0,966	0,559
Effet du positionnement de la sonde (LM : gabarit; TrA-OI-OE : cube de mousse)				
INTRA-évaluateur	0,030	0,013	0,734	0,713
INTER-évaluateurs	0,434	0,816	0,206	0,025

Les IDs et ESMs correspondant à l'utilisation d'une seule image ont été utilisés.

* Pour le LM, 8 valeurs (ID ou ESM) étaient comparées entre les tâches [LM-Bras vs LM-Jambe : 1 Variables (Contraction) × 2 Muscles × 2 Côtés × 2 Positions sonde = 8] et 16 valeurs étaient comparées entre les deux conditions de positionnement de la sonde [avec ou sans gabarit : 2 Variables (Repos et Contraction) × 2 Muscles × 2 Côtés × 2 Tâches = 16]).

† Pour TrA – OI et OE, 12 valeurs (ID ou ESM) étaient comparées entre les tâches [ÉAJT vs ÉAJF : 1 Variables (Contraction) × 3 Muscles × 2 Côtés × 2 Positions sonde = 12] et 24 valeurs étaient comparées entre les deux conditions de positionnement de la sonde [avec ou sans cube de mousse : 2 Variables (Repos et Contraction) × 3 Muscles × 2 Côtés × 2 Tâches = 24].

3.3.2.2 Effet de faire la moyenne de trois mesures par image sur la fidélité (muscles abdominaux)

Pour les muscles abdominaux, l'effet généré par la prise de trois mesures (une au centre et deux autres à un centimètre de chaque côté) pour chaque muscle sur la même image a été évalué. Cependant, il était tout d'abord nécessaire d'évaluer la faisabilité de cette pratique et de vérifier si ces mesures étaient comparables entre elles (bordures des muscles parallèles?). Il faut mentionner que les évaluateurs ne prenaient pas de mesures additionnelles à la mesure du centre, lorsque l'épaisseur du muscle n'était pas constante (déterminé par observation). Pour les tâches de ÉAJT et ÉAJF, les trois mesures ont été prises pour presque tous les muscles et sujets. Cependant, pour l'ADIM, c'était beaucoup plus difficile, car le mouvement latéral des muscles était beaucoup plus important, de sorte qu'une portion des muscles sortait du champ de vision de la sonde et qu'environ la moitié des mesures seulement ont pu être prises pour les trois muscles.

Des ANOVA à une voie pour mesures répétées, suivies du test de Tukey, ont été menées pour vérifier si des différences systématiques étaient présentes entre les trois mesures. Bien que les résultats fussent statistiquement significatifs dans plusieurs cas, les différences pouvaient être considérées comme petites (0,1 à 0,4 mm) selon le point de vue (**Tableau 8**). L'effet de faire la moyenne des trois mesures sur la fidélité a donc été évalué, tel que présenté à l'Annexe A, dans le **Tableau A-8**. Des analyses statistiques ont permis de comparer les ID et ESM (%) correspondant à une mesure comparativement à la moyenne de trois mesures (**Tableau 9**). Ces analyses ont permis de conclure à une amélioration de la fidélité INTRA- et INTER-évaluateurs,

ceci pour les deux indices de fidélité (ID et ESM) et pour les muscles TrA et OI. Cependant, bien que statistiquement significatifs, ces effets étaient petits, de l'ordre de 0,005 à 0,021 d'augmentation pour les IDs et de 0,4 à 1,7 % de diminution pour les ESMs (**Tableau 9**). Pour l'OE, seuls les ESMs de la fidélité INTRA-évaluateur se sont améliorées (de 0,8 %) alors que les IDs de la fidélité INTER-évaluateurs se sont détériorées (de 0,015). Il apparaît donc difficile de recommander de faire la moyenne de trois mesures, étant donné le peu de gain obtenu en termes de fidélité. Ces résultats peuvent être associés aux différences systématiques (biais) entre les trois mesures qui, bien qu'apparaissant relativement petites (ex. : 0,1 mm pour le TrA au repos), peuvent ne pas être négligeables, lorsque mises en relation avec le petit changement observé lors de la contraction. En effet, dans cet exemple, 0,1 mm représente 10 % du changement (1 mm) observé pour les tâches automatisées. Toujours pour le TrA, les mesures latérales étaient systématiquement plus élevées (3,1 mm) que les mesures localisées médialement (3,0 mm), ce qui suggère que cette stratégie de mesure pourrait être plus efficace si les trois mesures (ou plus) étaient prises où le muscle est plus uniforme, c'est-à-dire plus médialement. Cependant, les mêmes conclusions ne pourraient pas être appliquées pour l'OE où les différences entre les mesures étaient plus évidentes d'un côté à l'autre, ce qui complique la prise de mesures parallèles à travers les trois couches de muscles abdominaux. En somme, nous ne recommandons pas de faire la moyenne de mesures juxtaposées, du moins avec la méthodologie proposée ici.

Tableau 8 : Résultats statistiques correspondant à la comparaison entre les trois mesures effectuées pour chaque muscle abdominal, sur chaque image.

Mesure	Côté gauche		Côté droit	
	ANOVA (valeur P)	Tukey* (1 vs 2 vs 3)	ANOVA (valeur P)	Tukey* (1 vs 2 vs 3)
TrA - Repos	<0,001	1 > 2, 3 (3,1; 3,0; 3,0 mm)	0,001	1 > 2,3 (3,1; 3,0; 3,0 mm)
TrA - Contraction	0,267	/	0,601	/
OI - Repos	0,005	1 > 3 (7,0; 6,9; 6,8 mm)	0,077	/
OI - Contraction	0,481	/	0,614	/
OE - Repos	0,040	1 < 3 (5,3; 5,4; 5,5 mm)	<0,001	1 < 2,3; 2 < 3 (5,1; 5,3; 5,4 mm)
OE - Contraction	0,146	/	<0,001	1,2 < 3 (7,1; 7,3; 7,6 mm)

*La direction des différences entre les mesures 1, 2 et 3 est indiquée alors que les moyennes correspondant aux trois mesures sont identifiées entre parenthèses.

3.3.2.3 Implications pour la recherche et la clinique

Les résultats de l'étude D et les analyses statistiques réalisées avec les IDs et ESM (%) permettent de faire les recommandations suivantes :

- (1) Faire la moyenne de trois images;
- (2) Utiliser n'importe laquelle des tâches automatisées proposées (pour les deux groupes musculaires);
- (3) Ne pas utiliser le gabarit pour le LM, mais utiliser le cube de mousse pour les abdominaux;

- (4) Pour les abdominaux, ne faire qu'une seule mesure (centrale) par image;
- (5) Éviter d'utiliser différents évaluateurs, lorsque possible, pour comparer les mesures entre patients ou comparer les mesures d'un même patient.

En analysant les résultats correspondant à ces caractéristiques, il est possible de se prononcer sur la valeur de ces mesures sur le plan de la recherche et de la clinique. Premièrement, les IDs permettent de vérifier la capacité de ces tests à différencier les sujets entre eux. Il est recommandé d'exiger des CCIC (ou ID) supérieurs à 0,75 pour faire des comparaisons entre groupes (applications en recherche) et des valeurs supérieures à 0,90 pour le suivi individuel (applications cliniques) (Scientific Advisory Committee of the Medical Outcomes Trust 2002), bien que c'est la mesure ESM qui est la plus appropriée en ce qui a trait au suivi de patients. Selon ces recommandations, il apparaît que seules les mesures Repos et Contraction, et si on est un peu plus libéral, les mesures PTC, démontrent cette qualité à différencier les sujets entre eux (IDs > 0,75). Cependant, étant donné que les mesures de Repos et Contraction dépendent de la masse du sujet et du sexe (Rankin *et coll.* 2006), elles doivent être normalisées d'une quelconque façon, ce qui mène aux mesures dérivées qui sont moins fidèles. En somme, il apparaît que ces mesures ultrasons ne sont pas très utiles pour le diagnostic sur une base individuelle (application clinique). Cependant, la comparaison entre groupes est toujours possible, telle que généralement faite en recherche.

Tableau 9 : Résultats statistiques (valeurs P) correspondant aux Test-t (pour données paires) pour comparer les ID et les ESM correspondant à l'utilisation d'une ou de trois mesures (muscles abdominaux).

Type de fidélité	Indice de fidélité	1 mesure		3 mesures		Valeur P (Test-t)*
		Moyenne	Écart-type	Moyenne	Écart-type	
Muscle TrA						
INTRA-évaluateur	ID	0,837	0,042	0,853	0,041	0,009
	ESM (%)	12,1	2,3	11,4	2,5	0,002
INTER-évaluateurs	ID	0,780	0,044	0,801	0,044	0,027
	ESM (%)	16,2	2,7	15,0	2,9	0,001
Muscle OI						
INTRA-évaluateur	ID	0,915	0,027	0,920	0,024	0,014
	ESM (%)	9,2	1,5	8,8	1,4	0,002
INTER-évaluateurs	ID	0,866	0,039	0,880	0,035	<0,001
	ESM (%)	12,4	2,0	11,6	1,9	<0,001
Muscle OE						
INTRA-évaluateur	ID	0,834	0,067	0,844	0,062	0,117
	ESM (%)	12,5	1,3	11,7	2,1	0,001
INTER-évaluateurs	ID	0,810	0,110	0,795	0,120	0,015
	ESM (%)	14,0	3,2	14,2	3,5	0,301

Les IDs et ESMs correspondant à la moyenne de trois images ont été utilisés.

* Pour chaque Test-t, 16 valeurs (ID ou ESM) étaient comparées entre les deux façons d'analyser les muscles abdominaux [1 mesure (centrale) vs 3 mesures : 2 Variables (Repos et Contraction) × 2 Côtés × 2 Tâches × 2 Positions sonde = 16].

Les résultats des ESMs permettent de se prononcer sur la capacité à faire le suivi des patients (mesures répétées auprès du même individu). Dans ce contexte d'utilisation, plusieurs variables démontrent une excellente fidélité (ESM % < 10 %), notamment pour la fidélité INTRA-évaluateur et le muscle LM : Repos, Contraction, CRatio, SRRatio, SCRatio. Pour TrA, seules les variables Repos et PTC produisent d'aussi bons résultats (ESM % < 10 %). Bien entendu, la fidélité INTER-évaluateurs produit des résultats inférieurs, ce qui démontre l'importance de n'utiliser qu'un seul évaluateur pour faire le suivi d'un patient. Ici, contrairement aux applications visant à comparer des patients, il est possible d'utiliser des variables non normalisées (ex. : Repos, Contraction) pour faire le suivi des patients, étant donné que ceux-ci sont comparés à eux-mêmes. Des valeurs variant entre 10 et 20 % sont considérées comme typiques pour des mesures biologiques ou de performance, tel que discuté par Mannion et coll. (2008). Il est donc plus rare d'obtenir des résultats inférieurs à 10 %. La plus petite différence détectable ($PPDD_{95\%} = 1,96 \times \sqrt{2} \times ESM$), c.-à-d. la valeur au-delà de laquelle un effet de traitement n'est plus attribuable à une erreur de mesure, permet de mieux juger de l'importance d'une ESM de 10 %. La PPDD correspondant à une ESM de 10 % est de 28 %, ce qui est déjà considérable, car peu de traitements risquent de produire de tels effets thérapeutiques. Nous suggérons donc que des ESMs dépassant 10 % doivent être considérées avec beaucoup moins d'intérêt, notamment pour une application clinique (individuelle). Cependant, encore ici, il est toujours possible de comparer des groupes (application en recherche), mais il est évident que plus les ESMs seront élevées, plus il faudra de sujets pour obtenir une puissance statistique suffisante pour détecter des effets. Comme mentionné par Mannion et coll. (2008), de manière à déterminer si les erreurs sont acceptables en pratique, des études cliniques doivent être conduites afin de démontrer l'augmentation typique qui peut être attendue en ce qui a trait à l'épaisseur de ces muscles, ceci pour des interventions comme les exercices de stabilisation lombaire. Présentement, nous ne disposons pas de cette information pour donner des recommandations plus claires en relation avec les erreurs de mesure.

4. LIMITES

La présente étude comporte un nombre de sujets relativement important comparativement à la plupart des études de fidélité réalisées sur ces mesures (Costa *et coll.* 2009 ; Hebert *et coll.* 2009) et cette taille d'échantillon est en accord avec un nombre minimal variant entre 25 et 30 sujets de manière à obtenir des estimations plus robustes (Springate 2011). Cependant, ce nombre était insuffisant pour obtenir des estimations pour chaque groupe (sujets sains et lombalgiques) séparément, notamment pour les sujets lombalgiques qui représentent une population très hétérogène pour laquelle obtenir un échantillon représentatif demeure toujours un défi important. D'ailleurs, les sujets en phase aiguë n'ont pas été considérés, ce qui pourrait représenter une source de variabilité supplémentaire. De plus, la répartition des sujets selon le sexe était inégale entre les deux groupes. Il aurait été utile de comparer la fidélité des mesures provenant des deux groupes, mais il a été jugé plus utile de procurer des estimations de fidélité (ID et ESM) plus justes que de répondre à cette question. Le fait que la moitié de l'échantillon soit composé de sujets lombalgiques est d'ailleurs une force de cette étude, comparativement à ce qui avait été fait jusqu'à maintenant. Cela permet de mieux généraliser les résultats à la population pour laquelle ces mesures seront utilisées, bien qu'une évaluation spécifique aux sujets lombalgiques ait été préférable.

Le devis de recherche utilisé n'était pas complet, car seul l'évaluateur no 1 a fait les mesures sur deux jours différents. Il est possible que les résultats de fidélité intra-évaluateur aient été différents avec l'évaluateur no 2. Cela a empêché de considérer les trois sources d'erreur simultanément (évaluateur, jour, image) dans le modèle statistique. D'ailleurs, aucune étude n'a utilisé un tel devis de recherche pour ce type de mesures. Évidemment, il est question des ressources requises pour mener ce type d'étude. La présente étude se situe néanmoins relativement bien comparativement aux autres en termes du nombre de sujets, de l'inclusion de sujets lombalgiques chroniques, du type de fidélité évalué (INTER-évaluateurs en plus de l'INTRA-évaluateur) et du temps entre les sessions de mesures (INTRA-évaluateur sur des jours différents). Une seule autre étude a réuni toutes ces caractéristiques (Koppenhaver *et coll.* 2009), étude qui a d'ailleurs paru après que la demande de fonds pour le présent projet ait été soumise. La présente étude a permis en plus de répondre à des questions de recherche portant sur l'effet de la tâche et l'effet de mieux contrôler la position de la sonde sur la fidélité.

L'utilisation de ce type de mesures par ultrasons est présentement possible auprès des physiothérapeutes du Québec, mais dans un cadre relativement imprécis comprenant surtout la possibilité d'utiliser la rétroaction biologique (biofeedback) lors de l'entraînement de certains muscles profonds. Or, pour utiliser les mesures ultrasonographiques (telles que proposées dans ce rapport) dans le but de faire le suivi de la condition musculaire du patient durant le processus de réadaptation, il faudra une entente plus spécifique. Ce type d'application est présentement possible pour les physiothérapeutes d'autres pays et autres provinces canadiennes et cela fait présentement l'objet de nouvelles négociations au Québec. Il faut préciser que ces applications n'ont donc rien à voir avec le diagnostic. Il reste donc à définir les compétences associées à ces gestes, la formation nécessaire et les professions médicales pouvant les accomplir, ceci dans des contextes cliniques ou de recherche biomédicale.

5. CONCLUSION

La présente étude a mené à des mesures par ultrasons des muscles profonds du tronc comparables à ce qui a été publié ailleurs par d'autres groupes de recherche, démontrant ainsi le succès de l'implantation de cette technique de mesure par notre groupe de recherche. Nous disposons donc maintenant de l'expertise et des outils (interface d'analyse) nécessaires pour utiliser cette technique de mesure dans nos prochaines recherches. Il reste que la sensibilité au changement et l'utilité clinique de ce type de mesure, notamment en ce qui a trait à la relation de cause à effet entre les déficiences musculaires ainsi mesurées et la douleur ou les incapacités, restent à être établies avant de recommander leur utilisation en clinique.

Cette étude a aussi permis de documenter la fidélité INTRA- et INTER-évaluateurs de ces mesures chez un groupe de sujets, comprenant des sujets sains ainsi que des sujets lombalgiques chroniques, ce qui permet de mieux généraliser les résultats à la clientèle ciblée que par l'utilisation de sujets sains seulement, bien qu'un échantillon plus important de sujets lombalgiques aurait été préférable. Les résultats permettent tout de même de se prononcer sur les mesures (variables) à retenir et les mesures à rejeter pour une utilisation future, que ce soit pour des applications en recherche ou en clinique, bien que l'utilité de ce type de mesures de résultat reste à être démontrée avant de justifier un changement de pratique clinique.

Il a été démontré que différentes tâches automatisées peuvent produire des résultats comparables en termes de fidélité, ce qui laisse une plus grande marge de manœuvre aux futurs utilisateurs. Des recommandations sur les stratégies à utiliser ou non pour améliorer la fidélité, ceci sans trop investir de ressources, ont finalement été formulées.

6. IMPACT DU PROJET

Ce projet n'a pas fait l'objet d'un transfert de connaissances, car ces résultats ne se prêtent pas à ce genre d'exercice pour l'instant. En effet, bien que l'utilisation de l'imagerie par ultrasons soit plus accessible et populaire dans le milieu de la physiothérapie, les experts ne s'entendent pas sur son utilisation en clinique. Il est évident que cette technique de mesure peut être utilisée comme méthode de rétroaction biologique (biofeedback) pour aider le patient à mieux solliciter des muscles profonds, tels que le TrA (Henry and Teyhen 2007), notamment dans la phase initiale d'un programme d'exercices de stabilisation lombaire, tel qu'enseigné selon l'école de pensée australienne. Cependant, Stuart McGill (McGill 2002), pour sa part, ne met pas l'accent sur des muscles spécifiques comme, le TrA ou le LM, lors de l'exécution des exercices de stabilisation lombaire et ne requiert donc pas l'utilisation d'une méthode de *biofeedback* lors de la réadaptation. Conséquemment, utiliser cette technique en clinique pour produire des mesures de résultats (« outcome measure ») servant à documenter la progression des patients, telles qu'évaluées dans la présente étude, ne fait pas non plus l'unanimité. Les évidences scientifiques portant sur les effets réels de tels programmes d'exercices ne sont pas suffisamment claires pour justifier un tel investissement et de tels efforts en clinique. En effet, il n'est pas démontré que l'amélioration des patients est attribuable à une meilleure sollicitation des muscles spécifiques et profonds, tels que le LM et le TrA, même si une récente étude, bien que modeste (n = 40 patients), suggère que cela pourrait être le cas, mais pour le LM seulement (Hebert *et coll.* 2010). Nous croyons que ces mesures doivent être utilisées en recherche, afin de produire les évidences scientifiques nécessaires à un tel changement de pratique clinique, ce que nous projetons de réaliser prochainement à l'aide du protocole de mesures, du logiciel d'analyse et de certaines variables proposées dans ce document.

Suite des travaux. Ces tests peuvent être utilisés pour évaluer les fonctions neuromusculaires des muscles profonds du tronc. Ces tests ont le potentiel : (1) de suivre l'évolution des patients lors de la réadaptation, (2) de valider des modalités spécifiques d'intervention en réadaptation pour renverser les atteintes spécifiques identifiées et (3) d'aider à mieux se prononcer sur un retour au travail sécuritaire.

Les retombées éventuelles de ces tests neuromusculaires concernent, à court et moyen terme, la recherche portant sur la définition de groupes cliniques et sur l'étude des mécanismes d'action des interventions cliniques. Les études longitudinales requises pour étudier les mécanismes d'action des interventions cliniques permettront de documenter la sensibilité au changement de ces mesures. Elles concernent aussi, à plus long terme cette fois-ci, la clinique. En effet, considérant les difficultés à identifier un diagnostic chez les sujets lombalgiques chroniques, ces tests pourraient éventuellement être utilisés pour identifier des sous-groupes de patients, qui répondent bien ou pas du tout à certaines modalités spécifiques de traitement. Étant donné que diverses modalités de traitement sont prévues pour répondre à différentes pathologies spécifiques, il serait alors possible d'associer ce type de mesures par ultrasons à une condition clinique plus spécifique, donnant ainsi une certaine valeur diagnostique aux mesures utilisées. Les études portant sur le développement de règles de prédiction clinique (« clinical prediction rules ») suivent ce principe. Ces règles peuvent être appliquées à différents types d'intervention, passant des manipulations vertébrales (Childs *et coll.* 2004 ; Childs and Cleland 2006 ; Cleland *et coll.* 2007 ; Flynn *et coll.* 2002) aux exercices de stabilisation lombaire (Hicks *et coll.* 2005). Cette façon de faire permet d'attribuer des modalités d'intervention spécifiques à certains sous-

groupes de patients et d'obtenir des résultats cliniques plus importants (Brennan *et coll.* 2006 ; Childs *et coll.* 2004 ; Fritz *et coll.* 2003 ; Long *et coll.* 2004). Présentement, les mesures utilisées pour identifier ces sous-groupes sont de nature clinique (examen physique, questionnaires sur variables psychologiques). Il faut maintenant associer des mesures de nature physiologique à ces sous-groupes pour mieux comprendre les mécanismes d'action mis à contribution par ces interventions cliniques. Une fois ces mécanismes d'action identifiés, il sera possible de définir des modalités de traitement (programmes d'exercices) plus spécifiques. Ultimement, ces mesures pourraient même aider à raffiner les règles de prédiction clinique, tel que suggéré par une étude récente (Hebert *et coll.* 2010). Une telle démonstration permettrait de documenter la validité de construit de ce type de mesure et ainsi de mieux faire valoir leur utilité en clinique.

7. BIBLIOGRAPHIE

Airaksinen, O., Hildebrandt, J., Mannion, A. F., Brox, J. I., Klaber Moffett, J., Reis, S., Zanoli, G., Cedraschi, C., Kovacs, F. and Staal, J. B. European guidelines for the management of chronic non-specific low back pain. 1-91. 2004. Ref Type: Report

Baecke, J. A. H., Burema, J. and Frijters, J. E. R. (1982) A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *The American Journal of Clinical Nutrition*, **36**, 936-942.

Barker, K. L., Shamley, D. R. and Jackson, D. (2004) Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: the relationship to pain and disability. *Spine*, **29**, E515-E519.

Barr, K. P., Griggs, M. and Cadby, T. (2007) Lumbar stabilization: a review of core concepts and current literature, part 2. *Am.J.Phys.Med.Rehabil.*, **86**, 72-80.

Belavy, D. L., Armbrecht, G., Richardson, C. A., Felsenberg, D. and Hides, J. A. (2010) Muscle Atrophy and Changes in Spinal Morphology: Is the Lumbar Spine Vulnerable After Prolonged Bed-Rest? *Spine (Phila Pa 1976)*.

Borkan, J. M., Koes, B., Reis, S. and Cherkin, D. C. (1998) A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. *Spine*, **23**, 1992-1996.

Brennan, G. P., Fritz, J. M., Hunter, S. J., Thackeray, A., Delitto, A. and Erhard, R. E. (2006) Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. *Spine*, **31**, 623-631.

Childs, J. D. and Cleland, J. A. (2006) Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Phys.Ther.*, **86**, 122-131.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. and Delitto, A. (2004) A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann.Intern.Med.*, **141**, 920-928.

Cleland, J. A., Childs, J. D., Fritz, J. M., Whitman, J. M. and Eberhart, S. L. (2007) Development of a clinical prediction rule for guiding treatment of a subgroup of patients with neck pain: use of thoracic spine manipulation, exercise, and patient education. *Phys Ther*, **87**, 9-23.

Clinical research agenda for physical therapy (2000) *Phys Ther*, **80**, 499-513.

CLIP . Guide pratique : Clinique des lombalgies interdisciplinaire en première ligne (CLIP). 43 p. 2006. IRSST (Montréal), Direction de la santé publique (Montréal). Ref Type: Report

Costa, L. O., Maher, C. G., Latimer, J. and Smeets, R. J. (2009) Reproducibility of rehabilitative ultrasound imaging for the measurement of abdominal muscle activity: a systematic review. *Phys.Ther.*, **89**, 756-769.

CSST . Statistiques sur les affections vertébrales 2004-2007. 2008. Ref Type: Report

Dankaerts, W., O'Sullivan, P. B., Burnett, A. F., Straker, L. M. and Danneels, L. A. (2004) Reliability of EMG measurements for trunk muscles during maximal and sub-maximal voluntary isometric contractions in healthy controls and CLBP patients. *J.Electromyogr.Kinesiol.*, **14**, 333-342.

Danneels, L. A., Vanderstraeten, G. G., Cambier, D. C., Witvrouw, E. E. and De Cuyper, H. J. (2000) CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. *European Spine Journal*, **9**, 266-272.

Dionne, C. E. (1999) Low back pain. In Crombie, I. K. (ed), *Epidemiology of pain*. International Association for the Study of Pain Press, Seattle.

Durnin, J. V. G. A. and Womersley, J. (1974) Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years. *British Journal of Nutrition*, **32**, 77-97.

Ebenbichler, G. R., Oddsson, L. I., Kollmitzer, J. and Erim, Z. (2001) Sensory-motor control of the lower back: implications for rehabilitation. *Med.Sci.Sports Exerc.*, **33**, 1889-1898.

Eisenstein, S. M. (1999) "Instability" and low back pain. A way out of the semantic maze. In Szpalski, M., Gunzburg, R. and Pope, M. H. (eds), *Lumbar Segmental Instability*. Lippincott Williams & Wilkins, Philadelphia, pp. 39-51.

Ferreira, P. H., Ferreira, M. L. and Hodges, P. W. (2004) Changes in recruitment of the abdominal muscles in people with low back pain: ultrasound measurement of muscle activity. *Spine*, **29**, 2560-2566.

Ferreira, P. H., Ferreira, M. L., Maher, C. G., Refshauge, K., Herbert, R. D. and Hodges, P. W. (2010) Changes in recruitment of transversus abdominis correlate with disability in people with chronic low back pain. *Br.J Sports Med.*, **44**, 1166-1172.

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. and Allison, S. (2002) A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, **27**, 2835-2843.

Frank, J. W., Brooker, A. S., Demaio, S. E., Kerr, M. S., Maetzel, A., Shannon, H. S., Sullivan, T. J., Norman, R. W. and Wells, R. (1996) Disability resulting from occupational low back pain. Part II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. *Spine*, **21**, 2918-2929.

Fritz, J. M., Delitto, A. and Erhard, R. E. (2003) Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine*, **28**, 1363-1371.

Fritz, J. M., Erhard, R. E. and Hagen, B. F. (1998) Segmental instability of the lumbar spine. *Physical Therapy*, **78**, 889-896.

- Garrow, J. S. and Webster, J. (1985) Quetelet's index (W/H²) as a measure of fatness. *International Journal of Obesity*, **9**, 147-153.
- Hashemi, L., Webster, B. S., Edward, A. and Enerst, V. (1997) Length of Disability and Cost of Workers' Compensation Low Back Pain Claims. *Journal of Occupational and Environment Medicine*, **39**, 937-945.
- Hayden, J. A., van Tulder, M. W. and Tomlinson, G. (2005) Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann.Intern.Med.*, **142**, 776-785.
- Hebert, J. J., Koppenhaver, S. L., Magel, J. S. and Fritz, J. M. (2010) The relationship of transversus abdominis and lumbar multifidus activation and prognostic factors for clinical success with a stabilization exercise program: a cross-sectional study. *Arch.Phys.Med.Rehabil.*, **91**, 78-85.
- Hebert, J. J., Koppenhaver, S. L., Parent, E. C. and Fritz, J. M. (2009) A systematic review of the reliability of rehabilitative ultrasound imaging for the quantitative assessment of the abdominal and lumbar trunk muscles. *Spine (Phila Pa 1976)*, **34**, E848-E856.
- Henry, S. M. and Teyhen, D. S. (2007) Ultrasound imaging as a feedback tool in the rehabilitation of trunk muscle dysfunction for people with low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, **37**, 627-634.
- Henry, S. M. and Westervelt, K. C. (2005) The use of real-time ultrasound feedback in teaching abdominal hollowing exercises to healthy subjects. *J.Orthop.Sports Phys.Ther.*, **35**, 338-345.
- Henschke, N., Maher, C. G., Refshauge, K. M., Das, A. and McAuley, J. H. (2007) Low back pain research priorities: a survey of primary care practitioners. *BMC.Fam.Pract.*, **8**, 40.
- Hicks, G. E., Fritz, J. M., Delitto, A. and McGill, S. M. (2005) Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Arch.Phys Med.Rehabil.*, **86**, 1753-1762.
- Hides, J., Wilson, S., Stanton, W., McMahon, S., Keto, H., McMahon, K., Bryant, M. and Richardson, C. (2006) An MRI investigation into the function of the transversus abdominis muscle during "drawing-in" of the abdominal wall. *Spine*, **31**, E175-E178.
- Hides, J. A., Jull, G. A. and Richardson, C. A. (2001) Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine*, **26**, E243-E248.
- Hides, J. A., Miokovic, T., Belavy, D. L., Stanton, W. R. and Richardson, C. A. (2007) Ultrasound imaging assessment of abdominal muscle function during drawing-in of the abdominal wall: an intrarater reliability study. *Journal of Orthopaedic & Sports Physical Therapy*, **37**, 480-486.

- Hides, J. A., Richardson, C. A. and Jull, G. A. (1995) Magnetic resonance imaging and ultrasonography of the lumbar multifidus muscle. Comparison of two different modalities. *Spine*, **20**, 54-58.
- Hides, J. A., Stokes, M. J., Saide, M., Jull, G. A. and Cooper, D. H. (1994) Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain. *Spine*, **19**, 165-172.
- Hildebrandt, V. H. (1995) Back pain in the working population: prevalence rates in Dutch trades and professions. *Ergonomics*, **38**, 1283-1298.
- Hodges, P. W. and Moseley, G. L. (2003) Pain and motor control of the lumbopelvic region: effect and possible mechanisms. *J.Electromyogr.Kinesiol.*, **13**, 361-370.
- Hodges, P. W. and Richardson, C. A. (1999) Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. *Archives of Physical Medicine and Rehabilitation*, **80**, 1005-1012.
- Hyun, J. K., Lee, J. Y., Lee, S. J. and Jeon, J. Y. (2007) Asymmetric atrophy of multifidus muscle in patients with unilateral lumbosacral radiculopathy. *Spine*, **32**, E598-E602.
- Kiesel, K. B., Uhl, T. L., Underwood, F. B., Rodd, D. W. and Nitz, A. J. (2007a) Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. *Man.Ther.*, **12**, 161-166.
- Kiesel, K. B., Underwood, F. B., Maticolla, C., Nitz, A. J. and Malone, T. R. (2007b) A comparison of select trunk muscle thickness change between subjects with low back pain classified in the treatment-based classification system and asymptomatic controls. *J.Orthop.Sports Phys.Ther.*, **37**, 596-607.
- Koppenhaver, S. L., Hebert, J. J., Fritz, J. M., Parent, E. C., Teyhen, D. S. and Magel, J. S. (2009) Reliability of rehabilitative ultrasound imaging of the transversus abdominis and lumbar multifidus muscles. *Arch.Phys.Med.Rehabil.*, **90**, 87-94.
- Langevin, H. M., Stevens-Tuttle, D., Fox, J. R., Badger, G. J., Bouffard, N. A., Krag, M. H., Wu, J. and Henry, S. M. (2009) Ultrasound evidence of altered lumbar connective tissue structure in human subjects with chronic low back pain. *BMC.Musculoskelet.Disord.*, **10**, 151.
- Leboeuf-Yde, C. and Manniche, C. (2001) Low back pain: Time to get off the treadmill. *J Manipulative Physiol Ther*, **24**, 63-66.
- Long, A., Donelson, R. and Fung, T. (2004) Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine*, **29**, 2593-2602.
- MacDonald, D., Moseley, G. L. and Hodges, P. W. (2009) Why do some patients keep hurting their back? Evidence of ongoing back muscle dysfunction during remission from recurrent back pain. *Pain*, **142**, 183-188.

- Mannion, A. F., Pulkovski, N., Gubler, D., Gorelick, M., O'Riordan, D., Loupas, T., Schenk, P., Gerber, H. and Sprott, H. (2008) Muscle thickness changes during abdominal hollowing: an assessment of between-day measurement error in controls and patients with chronic low back pain. *Eur.Spine J.*, **17**, 494-501.
- Marras, W. S., Jorgensen, M. J., Granata, K. P. and Wiand, B. (2001) Female and male trunk geometry: size and prediction of the spine loading trunk muscles derived from MRI. *Clinical Biomechanics*, **16**, 38-46.
- McGill, S. (2002) *Low back disorders. Evidence-based prevention and rehabilitation*. Human Kinetics, Champaign, U.S.A.
- McMeeken, J. M., Beith, I. D., Newham, D. J., Milligan, P. and Critchley, D. J. (2004) The relationship between EMG and change in thickness of transversus abdominis. *Clin Biomech (Bristol., Avon.)*, **19**, 337-342.
- Moseley, G. and Hodges, P. (2005) Chronic pain and motor control. In Jull, G. and Boyling, J. (eds), *Grievous modern manual therapy of the vertebral column*. Churchill-Livingstone, Edinburgh, pp. 215-232.
- Norasteh, A., Ebrahimi, E., Salavati, M., Rafiei, J. and Abbasnejad, E. Reliability of B-mode ultrasonography for abdominal muscles in asymptomatic and patients with acute low back pain. *Journal of bodywork and movement therapies* **11**, 17-20. 2007. Ref Type: Journal (Full)
- O'Sullivan, P. B., Twomey, L. and Allison, G. T. (1998) Altered abdominal muscle recruitment in patients with chronic back pain following a specific exercise intervention. *Journal of Orthopaedic & Sports Physical Therapy*, **27**, 114-124.
- O'Sullivan, P. B., Twomey, L. T. and Allison, G. T. (1997) Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine*, **22**, 2959-2967.
- Panjabi, M. M. (2003) Clinical spinal instability and low back pain. *J.Electromyogr.Kinesiol.*, **13**, 371-379.
- Panjabi, M. M. (2006) A hypothesis of chronic back pain: ligament subfailure injuries lead to muscle control dysfunction. *Eur Spine J*, **15**, 668-676.
- Pruimboom, L. and van Dam, A. C. (2007) Chronic pain: a non-use disease. *Med.Hypotheses*, **68**, 506-511.
- Rankin, G., Stokes, M. and Newham, D. J. (2006) Abdominal muscle size and symmetry in normal subjects. *Muscle Nerve*, **34**, 320-326.
- Reeves, N. P., Narendra, K. S. and Cholewicki, J. (2007) Spine stability: the six blind men and the elephant. *Clin.Biomech.(Bristol., Avon.)*, **22**, 266-274.

Richardson, C., Hodges, P. and Hides, J. (2004) *Therapeutic exercise for lumbopelvic stabilization*. Churchill Livingstone, Sydney.

Scientific Advisory Committee of the Medical Outcomes Trust (2002) Assessing health status and quality-of-life instruments: Attributes and review criteria. *Qual.Life Res*, **11**, 193-205.

Shavelson, R. J. and Webb, N. M. (1991) *Generalizability theory. A primer*. SAGE Publications, London.

Shrout, P. E. and Fleiss, J. L. (1979) Intraclass Correlations: Uses in assessing rater reliability. *Psychological Bulletin*, **86**, 420-428.

Smeets, R. J., Vlaeyen, J. W., Kester, A. D. and Knottnerus, J. A. (2006) Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *J Pain*, **7**, 261-271.

Smidt, N., de Vet, H. C., Bouter, L. M., Dekker, J., Arendzen, J. H., De Bie, R. A., Bierma-Zeinstra, S. M., Helders, P. J., Keus, S. H., Kwakkel, G., Lenssen, T., Oostendorp, R. A., Ostelo, R. W., Reijman, M., Terwee, C. B., Theunissen, C., Thomas, S., van Baar, M. E., van 't, H. A., van Peppen, R. P., Verhagen, A. and van der Windt, D. A. (2005) Effectiveness of exercise therapy: a best-evidence summary of systematic reviews. *Aust J Physiother.*, **51**, 71-85.

Spitzer, W. O., Leblanc, F. E. and Dupuis, M. (1987) Scientific approach to the assessment and management of activity-related spinal disorders. *Spine*, **12**, S1-S59.

Springate, S. D. (2011) The effect of sample size and bias on the reliability of estimates of error: a comparative study of Dahlberg's formula. *Eur J Orthod.*, doi:[10.1093/ejo/cjr010](https://doi.org/10.1093/ejo/cjr010).

Stokes, M., Hides, J., Elliott, J., Kiesel, K. and Hodges, P. (2007) Rehabilitative ultrasound imaging of the posterior paraspinal muscles. *Journal of Orthopaedic & Sports Physical Therapy*, **37**, 581-595.

Stokes, M., Rankin, G. and Newham, D. J. (2005) Ultrasound imaging of lumbar multifidus muscle: normal reference ranges for measurements and practical guidance on the technique. *Man.Ther.*, **10**, 116-126.

Teyhen, D. S., Gill, N. W., Whittaker, J. L., Henry, S. M., Hides, J. A. and Hodges, P. (2007) Rehabilitative ultrasound imaging of the abdominal muscles. *Journal of Orthopaedic & Sports Physical Therapy*, **37**, 450-466.

Teyhen, D. S., Miltenberger, C. E., Deiters, H. M., Del Toro, Y. M., Pulliam, J. N., Childs, J. D., Boyles, R. E. and Flynn, T. W. (2005) The use of ultrasound imaging of the abdominal drawing-in maneuver in subjects with low back pain. *J.Orthop.Sports Phys.Ther.*, **35**, 346-355.

Teyhen, D. S., Williamson, J. N., Carlson, N. H., Suttles, S. T., O'Laughlin, S. J., Whittaker, J. L., Goffar, S. L. and Childs, J. D. (2009) Ultrasound characteristics of the deep abdominal muscles during the active straight leg raise test. *Arch.Phys.Med.Rehabil.*, **90**, 761-767.

van Tulder, M., Malmivaara, A., Esmail, R. and Koes, B. (2000) Exercise therapy for low back pain. A systematic review within the framework of the Cochrane collaboration back review group. *Spine*, **25**, 2784-2796.

Vasseljen, O. and Fladmark, A. M. (2010) Abdominal muscle contraction thickness and function after specific and general exercises: A randomized controlled trial in chronic low back pain patients. *Man.Ther.*, **15**, 482-489.

Vuori, I. M. (2001) Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. *Med.Sci.Sports Exerc.*, **33**, S551-S586.

Wallwork, T. L., Hides, J. A. and Stanton, W. R. (2007) Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging. *Journal of Orthopaedic & Sports Physical Therapy*, **37**, 608-612.

Wallwork, T. L., Stanton, W. R., Freke, M. and Hides, J. A. (2008) The effect of chronic low back pain on size and contraction of the lumbar multifidus muscle. *Man.Ther.*.

Watanabe, K., Miyamoto, K., Masuda, T. and Shimizu, K. (2004) Use of ultrasonography to evaluate thickness of the erector spinae muscle in maximum flexion and extension of the lumbar spine. *Spine*, **29**, 1472-1477.

Wessels, T., van Tulder, M., Sigl, T., Ewert, T., Limm, H. and Stucki, G. (2006) What predicts outcome in non-operative treatments of chronic low back pain? A systematic review. *Eur Spine J*, **15**, 1633-1644.

Whittaker, J. L., Teyhen, D. S., Elliott, J. M., Cook, K., Langevin, H. M., Dahl, H. H. and Stokes, M. (2007) Rehabilitative ultrasound imaging: understanding the technology and its applications. *J.Orthop.Sports Phys.Ther.*, **37**, 434-449.

Whittaker, J. L., Warner, M. B. and Stokes, M. J. (2009) Induced transducer orientation during ultrasound imaging: effects on abdominal muscle thickness and bladder position. *Ultrasound Med.Biol.*, **35**, 1803-1811.

ANNEXE A : RÉSULTATS DE FIDÉLITÉ (ÉTUDE D)

Tableau A-1 : Résultats de fidélité (étude D : 1 image) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures Repos, Contraction et DR des muscles LM (à L4) et TrA du côté gauche

Mesure	Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
Repos	LM	Bras	N	32,9	6,0	0,89	2,1	6	30,7	6,2	0,63	3,9	13
			O	33,3	5,8	0,87	2,1	6	31,6	6,6	0,59	4,3	14
		Jambe	N	33,7	6,2	0,88	2,2	7	31,4	6,2	0,65	3,9	13
			O	33,7	6,4	0,89	2,2	6	31,9	6,8	0,63	4,4	14
	TrA	ÉAJT	N	3,1	0,9	0,81	0,4	13	3,0	0,9	0,74	0,5	16
			O	3,1	0,8	0,85	0,3	10	3,0	1,0	0,77	0,5	15
		ÉAJF	N	3,1	0,9	0,79	0,4	13	3,0	0,9	0,77	0,4	15
			O	3,1	0,8	0,82	0,3	11	2,9	0,9	0,68	0,5	17
		ADIM	O	3,2	0,9	0,91	0,3	8	3,1	0,8	0,59	0,5	18
		Contraction	LM	Bras	N	39,7	6,6	0,90	2,1	5	37,7	7,1	0,74
O	39,9				5,9	0,94	1,5	4	38,3	6,9	0,73	3,7	10
Jambe	N			41,4	6,7	0,91	2,0	5	39,1	6,8	0,71	3,9	10
	O			41,5	6,3	0,92	1,8	4	39,6	7,0	0,63	4,4	11
TrA	ÉAJT		N	3,9	1,2	0,74	0,6	16	3,8	1,4	0,70	0,8	21
			O	3,9	1,2	0,80	0,5	13	3,7	1,3	0,76	0,6	18
	ÉAJF		N	4,5	1,5	0,77	0,7	17	4,2	1,6	0,78	0,8	19
			O	4,5	1,5	0,88	0,5	12	4,1	1,6	0,73	0,9	21
	ADIM		O	6,1	1,9	0,88	0,7	11	6,0	2,1	0,79	1,0	16
	DiffRel		LM	Bras	N	21,5	10,1	0,57	6,8	32	24,0	13,6	0,28
O		20,8			10,0	0,43	7,6	37	22,5	12,7	0,06	12,5	55
Jambe		N		23,6	9,3	0,42	7,3	31	25,8	12,6	0,36	10,4	40
		O		23,9	9,4	0,49	6,9	29	25,4	11,8	0,45	9,3	36
TrA		ÉAJT	N	27,5	31,5	0,53	21,8	79	29,4	33,9	0,34	28,9	98
			O	27,0	32,9	0,53	22,9	85	23,5	29,5	0,52	20,6	88
		ÉAJF	N	46,5	36,3	0,50	27,2	59	39,6	36,4	0,48	29,8	75
			O	46,0	39,3	0,71	21,4	47	39,9	40,7	0,59	26,4	66
		ADIM	O	95,2	47,1	0,71	26,3	28	98,3	64,4	0,66	39,8	40

LM : multifides lombaires; TrA : transverse de l'abdomen; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M). Les ID $\geq 0,75$ et ESM (%) $\leq 10\%$ (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-2 : Résultats de fidélité (étude D : 1 image) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures CRatio, SRRatio, SCRatio des muscles LM (à L4) et TrA du côté gauche

Mesure	Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
CRatio	LM	Bras	N	1,22	0,10	0,57	0,07	6	1,24	0,14	0,28	0,12	10
			O	1,21	0,10	0,43	0,08	6	1,23	0,13	0,06	0,12	10
		Jambe	N	1,24	0,09	0,42	0,07	6	1,26	0,13	0,36	0,10	8
			O	1,24	0,09	0,49	0,07	6	1,25	0,12	0,45	0,09	7
	TrA	ÉAJT	N	1,28	0,32	0,53	0,22	17	1,29	0,34	0,34	0,29	22
			O	1,27	0,33	0,53	0,23	18	1,24	0,29	0,52	0,21	17
		ÉAJF	N	1,46	0,36	0,50	0,27	19	1,40	0,36	0,48	0,30	21
			O	1,46	0,39	0,71	0,21	15	1,40	0,41	0,59	0,26	19
		ADIM	O	1,95	0,47	0,71	0,26	13	1,98	0,64	0,66	0,40	20
		SRRatio	LM	Bras	N	1,03	0,10	0,30	0,08	8	1,04	0,20	0,18
O	1,05				0,09	0,16	0,09	8	1,05	0,13	0,27	0,12	11
Jambe	N			1,04	0,09	0,21	0,09	8	1,05	0,16	0,16	0,15	14
	O			1,04	0,09	0,22	0,08	8	1,02	0,12	0,09	0,12	12
TrA	ÉAJT		N	1,02	0,16	0,27	0,14	14	1,02	0,24	0,25	0,23	22
			O	1,06	0,17	0,25	0,15	14	1,06	0,21	0,31	0,17	17
	ÉAJF		N	1,03	0,18	0,28	0,15	15	1,05	0,25	0,36	0,22	21
			O	1,03	0,18	0,25	0,16	16	1,07	0,23	0,00	0,24	22
	ADIM		O	1,04	0,12	0,05	0,12	12	1,06	0,20	0,00	0,22	21
	SCRatio		LM	Bras	N	1,01	0,08	0,36	0,06	6	1,03	0,14	0,29
O		1,02			0,07	0,51	0,06	5	1,02	0,11	0,38	0,09	9
Jambe		N		1,02	0,08	0,29	0,07	6	1,03	0,14	0,16	0,13	13
		O		1,02	0,07	0,28	0,06	6	1,01	0,10	0,23	0,09	9
TrA		ÉAJT	N	1,05	0,24	0,36	0,20	19	1,10	0,36	0,30	0,32	29
			O	1,09	0,28	0,52	0,20	18	1,10	0,31	0,33	0,27	25
		ÉAJF	N	1,08	0,24	0,30	0,22	20	1,11	0,33	0,23	0,31	28
			O	1,10	0,24	0,51	0,17	15	1,09	0,30	0,11	0,29	26
		ADIM	O	1,01	0,15	0,00	0,19	19	1,03	0,22	0,23	0,23	22

LM : multifides lombaires; TrA : transverse de l'abdomen; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M). Les ID $\geq 0,75$ et ESM (%) $\leq 10\%$ (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-3 : Résultats de fidélité (étude D : 1 image) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures PTC (pourcentage de l'épaisseur de tous les abdominaux lors de la contraction) des muscles abdominaux (TrA, OI et OE) du côté gauche

Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
			M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
TrA	ÉAJT	N	21,2	4,6	0,73	2,4	11	20,8	4,7	0,59	3,1	15
		O	20,9	4,6	0,55	3,2	15	20,0	5,4	0,60	3,5	17
	ÉAJF	N	21,2	5,2	0,70	2,9	14	20,1	4,9	0,70	2,8	14
		O	21,3	4,7	0,78	2,2	10	20,1	4,8	0,53	3,4	17
ADIM	O	27,3	3,8	0,61	2,5	9	28,1	5,8	0,38	4,8	17	
OI	ÉAJT	N	41,4	5,3	0,71	2,9	7	42,6	5,3	0,59	3,6	8
		O	42,0	5,6	0,71	3,0	7	41,7	5,9	0,51	4,2	10
	ÉAJF	N	40,6	5,7	0,65	3,5	9	41,6	6,2	0,55	4,3	10
		O	41,1	5,9	0,71	3,2	8	41,3	6,1	0,63	3,8	9
ADIM	O	47,1	5,4	0,62	3,4	7	46,5	5,7	0,07	5,9	13	
OE	ÉAJT	N	37,4	6,0	0,81	2,7	7	36,6	6,6	0,60	4,2	12
		O	37,1	6,5	0,79	3,0	8	38,3	7,5	0,60	4,9	13
	ÉAJF	N	38,2	6,8	0,60	4,4	11	38,3	7,7	0,59	5,1	13
		O	37,5	7,5	0,82	3,2	9	38,6	7,9	0,69	4,5	12
ADIM	O	25,6	6,4	0,71	3,6	14	25,5	5,7	0,40	4,7	18	

TrA : transverse de l'abdomen; OI : oblique interne; OE : oblique externe; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M). Les ID \geq 0,75 et ESM (%) \leq 10 % (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-4 : Résultats de fidélité (étude D : 3 images) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures Repos, Contraction et DiffRel des muscles LM (à L4) et TrA du côté gauche

Mesure	Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
Repos	LM	Bras	N	32,9	6,0	0,92	1,7	5	30,7	6,2	0,68	3,5	11
			O	33,3	5,8	0,90	1,8	5	31,6	6,6	0,62	4,1	13
		Jambe	N	33,7	6,2	0,91	1,9	6	31,4	6,2	0,68	3,6	12
			O	33,7	6,4	0,91	1,9	6	31,9	6,8	0,65	4,2	13
	TrA	ÉAJT	N	3,1	0,9	0,86	0,3	11	3,0	0,9	0,82	0,4	13
			O	3,1	0,8	0,89	0,3	9	3,0	1,0	0,83	0,4	13
		ÉAJF	N	3,1	0,9	0,87	0,3	10	3,0	0,9	0,85	0,3	12
			O	3,1	0,8	0,90	0,2	8	2,9	0,9	0,77	0,4	13
		ADIM	O	3,2	0,9	0,94	0,2	7	3,1	0,8	0,64	0,5	16
Contraction	LM	Bras	N	39,7	6,6	0,93	1,7	4	37,7	7,1	0,83	2,9	8
			O	39,9	5,9	0,96	1,1	3	38,3	6,9	0,76	3,4	9
		Jambe	N	41,4	6,7	0,94	1,7	4	39,1	6,8	0,74	3,6	9
			O	41,5	6,3	0,94	1,5	4	39,6	7,0	0,66	4,1	10
	TrA	ÉAJT	N	3,9	1,2	0,77	0,6	15	3,8	1,4	0,76	0,7	18
			O	3,9	1,2	0,87	0,4	11	3,7	1,3	0,82	0,5	15
		ÉAJF	N	4,5	1,5	0,82	0,7	15	4,2	1,6	0,85	0,6	15
			O	4,5	1,5	0,92	0,4	10	4,1	1,6	0,80	0,7	18
		ADIM	O	6,1	1,9	0,90	0,6	10	6,0	2,1	0,82	0,9	15
DiffRel	LM	Bras	N	21,5	10,1	0,71	5,0	23	24,0	13,6	0,40	9,0	37
			O	20,8	10,0	0,57	5,8	28	22,5	12,7	0,09	10,3	46
		Jambe	N	23,6	9,3	0,54	5,6	24	25,8	12,6	0,44	8,7	34
			O	23,9	9,4	0,67	4,8	20	25,4	11,8	0,57	7,2	29
	TrA	ÉAJT	N	27,5	31,5	0,62	18,1	66	29,4	33,9	0,47	21,9	75
			O	27,0	32,9	0,62	18,9	70	23,5	29,5	0,68	14,9	63
		ÉAJF	N	46,5	36,3	0,64	20,5	44	39,6	36,4	0,69	19,1	48
			O	46,0	39,3	0,84	14,8	32	39,9	40,7	0,73	19,1	48
		ADIM	O	95,2	47,1	0,82	19,2	20	98,3	64,4	0,76	30,8	31

LM : multifides lombaires; TrA : transverse de l'abdomen; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M). Les ID $\geq 0,75$ et ESM (%) $\leq 10\%$ (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-5 : Résultats de fidélité (étude D : 3 images) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures CRatio, SRRatio, SCRatio des muscles LM (à L4) et TrA du côté gauche

Mesure	Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
CRatio	LM	Bras	N	1,22	0,10	0,71	0,05	4	1,24	0,14	0,40	0,09	7
			O	1,21	0,10	0,57	0,06	5	1,23	0,13	0,09	0,10	8
		Jambe	N	1,24	0,09	0,54	0,06	5	1,26	0,13	0,44	0,09	7
			O	1,24	0,09	0,67	0,05	4	1,25	0,12	0,57	0,07	6
	TrA	ÉAJT	N	1,28	0,32	0,62	0,18	14	1,29	0,34	0,47	0,22	17
			O	1,27	0,33	0,62	0,19	15	1,24	0,29	0,68	0,15	12
		ÉAJF	N	1,46	0,36	0,64	0,21	14	1,40	0,36	0,69	0,19	14
			O	1,46	0,39	0,84	0,15	10	1,40	0,41	0,73	0,19	14
		ADIM	O	1,95	0,47	0,82	0,19	10	1,98	0,64	0,76	0,31	16
SRRatio	LM	Bras	N	1,03	0,10	0,48	0,06	6	1,04	0,20	0,23	0,15	15
			O	1,05	0,09	0,20	0,07	7	1,05	0,13	0,33	0,10	10
		Jambe	N	1,04	0,09	0,28	0,07	7	1,05	0,16	0,21	0,13	12
			O	1,04	0,09	0,31	0,07	6	1,02	0,12	0,14	0,10	10
	TrA	ÉAJT	N	1,02	0,16	0,35	0,12	12	1,02	0,24	0,34	0,18	18
			O	1,06	0,17	0,31	0,13	12	1,06	0,21	0,45	0,13	12
		ÉAJF	N	1,03	0,18	0,38	0,12	12	1,05	0,25	0,62	0,13	13
			O	1,03	0,18	0,40	0,12	11	1,07	0,23	0,00	0,19	18
		ADIM	O	1,04	0,12	0,08	0,10	9	1,06	0,20	0,00	0,18	17
SCRatio	LM	Bras	N	1,01	0,08	0,50	0,05	5	1,03	0,14	0,38	0,09	9
			O	1,02	0,07	0,65	0,04	4	1,02	0,11	0,46	0,07	7
		Jambe	N	1,02	0,08	0,35	0,06	5	1,03	0,14	0,21	0,11	11
			O	1,02	0,07	0,35	0,05	5	1,01	0,10	0,33	0,07	7
	TrA	ÉAJT	N	1,05	0,24	0,44	0,17	16	1,10	0,36	0,46	0,23	21
			O	1,09	0,28	0,65	0,15	14	1,10	0,31	0,47	0,20	19
		ÉAJF	N	1,08	0,24	0,46	0,16	15	1,11	0,33	0,31	0,25	23
			O	1,10	0,24	0,64	0,13	12	1,09	0,30	0,16	0,24	22
		ADIM	O	1,01	0,15	0,00	0,14	14	1,03	0,22	0,39	0,16	15

LM : multifides lombaires; TrA : transverse de l'abdomen; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M), Les ID $\geq 0,75$ et ESM (%) $\leq 10\%$ (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-6 : Résultats de fidélité (étude D : 3 images) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures PTC (pourcentage de l'épaisseur totale de tous les abdominaux lors de la contraction) des muscles abdominaux (TrA, OI et OE) du côté gauche

Muscle	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
			M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
TrA	ÉAJT	N	21,2	4,6	0,79	2,0	10	20,8	4,7	0,66	2,6	13
		O	20,9	4,6	0,65	2,6	13	20,0	5,4	0,67	3,0	15
	ÉAJF	N	21,2	5,2	0,75	2,6	12	20,1	4,9	0,79	2,2	11
		O	21,3	4,7	0,86	1,7	8	20,1	4,8	0,61	2,9	14
	ADIM	O	27,3	3,8	0,72	1,9	7	28,1	5,8	0,43	4,3	15
OI	ÉAJT	N	41,4	5,3	0,76	2,5	6	42,6	5,3	0,67	3,0	7
		O	42,0	5,6	0,76	2,7	6	41,7	5,9	0,55	3,8	9
	ÉAJF	N	40,6	5,7	0,73	2,9	7	41,6	6,2	0,63	3,6	9
		O	41,1	5,9	0,78	2,7	7	41,3	6,1	0,68	3,4	8
	ADIM	O	47,1	5,4	0,72	2,7	6	46,5	5,7	0,08	5,2	11
OE	ÉAJT	N	37,4	6,0	0,86	2,2	6	36,6	6,6	0,70	3,4	9
		O	37,1	6,5	0,83	2,6	7	38,3	7,5	0,66	4,3	11
	ÉAJF	N	38,2	6,8	0,66	3,8	10	38,3	7,7	0,68	4,2	11
		O	37,5	7,5	0,87	2,7	7	38,6	7,9	0,74	4,0	10
	ADIM	O	25,6	6,4	0,80	2,8	11	25,5	5,7	0,50	3,8	15

TrA : transverse de l'abdomen; OI : oblique interne; OE : oblique externe; ÉAJT : élévation active jambe tendue; ÉAJF : élévation active jambe fléchie; ADIM : « abdominal drawing-in maneuver »; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M), Les ID $\geq 0,75$ et ESM (%) ≤ 10 % (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-7 : Résultats de fidélité (étude D : 1 ou 3 images) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures Repos, Contraction et CRatio du muscle LM à L2 du côté gauche

Mesure	Étude D	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
Repos	1 image	Bras	N	27,7	5,3	0,84	2,2	8	26,8	5,8	0,78	2,8	10
			O	27,4	5,0	0,86	1,9	7	27,2	5,8	0,70	3,3	12
		Jambe	N	28,2	5,5	0,83	2,3	8	27,0	5,7	0,81	2,6	10
			O	27,8	5,5	0,91	1,7	6	27,0	5,9	0,75	3,0	11
	3 images	Bras	N	27,7	5,3	0,90	1,7	6	26,8	5,8	0,84	2,3	8
			O	27,4	5,0	0,92	1,4	5	27,2	5,8	0,77	2,7	10
		Jambe	N	28,2	5,5	0,88	1,9	7	27,0	5,7	0,86	2,1	8
			O	27,8	5,5	0,95	1,3	5	27,0	5,9	0,80	2,6	10
Contraction	1 image	Bras	N	35,4	6,7	0,88	2,4	7	34,3	6,9	0,83	2,9	9
			O	34,9	6,5	0,91	2,0	6	34,7	7,3	0,87	2,7	8
		Jambe	N	35,5	6,6	0,93	1,8	5	33,8	7,4	0,74	3,9	12
			O	35,3	6,5	0,93	1,8	5	34,3	6,7	0,84	2,7	8
	3 images	Bras	N	35,4	6,7	0,92	1,8	5	34,3	6,9	0,89	2,3	7
			O	34,9	6,5	0,93	1,7	5	34,7	7,3	0,91	2,2	6
		Jambe	N	35,5	6,6	0,95	1,4	4	33,8	7,4	0,76	3,7	11
			O	35,3	6,5	0,96	1,3	4	34,3	6,7	0,87	2,4	7
CRatio	1 image	Bras	N	1,3	0,1	0,44	0,1	8	1,3	0,2	0,24	0,2	12
			O	1,3	0,1	0,41	0,1	7	1,3	0,1	0,26	0,1	10
		Jambe	N	1,3	0,1	0,19	0,1	8	1,3	0,1	0,22	0,1	11
			O	1,3	0,1	0,51	0,1	6	1,3	0,2	0,17	0,1	11
	3 images	Bras	N	1,3	0,1	0,62	0,1	5	1,3	0,2	0,48	0,1	7
			O	1,3	0,1	0,57	0,1	5	1,3	0,1	0,39	0,1	7
		Jambe	N	1,3	0,1	0,26	0,1	6	1,3	0,1	0,30	0,1	9
			O	1,3	0,1	0,71	0,0	4	1,3	0,2	0,22	0,1	10

LM : multifides lombaires; PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M), Les ID $\geq 0,75$ et ESM (%) ≤ 10 % (critères d'excellente fidélité) sont identifiés en caractères gras.

Tableau A-8 : Résultats de fidélité (étude D : 3 images) INTRA-évaluateur et INTER-évaluateurs correspondant aux mesures Repos, Contraction et CRatio du muscle TrA du côté gauche : comparaison entre 1 et 3 mesures par image

Mesure	Étude D	Tâche	PS	Fidélité INTRA-évaluateur					Fidélité INTER-évaluateurs				
				M	ET	ID	ESM	ESM (%)	M	ET	ID	ESM	ESM (%)
Repos	1 mesure	ÉAJT	N	3,1	0,9	0,86	0,3	11	3,0	0,9	0,82	0,4	13
			O	3,1	0,8	0,89	0,3	9	3,0	1,0	0,83	0,4	13
		ÉAJF	N	3,1	0,9	0,87	0,3	10	3,0	0,9	0,85	0,3	12
			O	3,1	0,8	0,90	0,2	8	2,9	0,9	0,77	0,4	13
	3 mesures	ÉAJT	N	3,1	0,9	0,88	0,3	10	3,0	0,9	0,78	0,4	14
			O	3,1	0,8	0,92	0,2	7	3,0	0,9	0,84	0,4	12
		ÉAJF	N	3,1	0,9	0,88	0,3	9	3,0	0,9	0,85	0,3	12
			O	3,1	0,8	0,89	0,2	8	3,0	0,8	0,82	0,3	12
Contraction	1 mesure	ÉAJT	N	3,9	1,2	0,77	0,6	15	3,8	1,4	0,76	0,7	18
			O	3,9	1,2	0,87	0,4	11	3,7	1,3	0,82	0,5	15
		ÉAJF	N	4,5	1,5	0,82	0,7	15	4,2	1,6	0,85	0,6	15
			O	4,5	1,5	0,92	0,4	10	4,1	1,6	0,80	0,7	18
	3 mesures	ÉAJT	N	3,8	1,2	0,83	0,5	13	3,7	1,4	0,80	0,6	16
			O	3,9	1,1	0,86	0,4	11	3,7	1,2	0,84	0,5	13
		ÉAJF	N	4,4	1,5	0,84	0,6	14	4,1	1,5	0,82	0,6	15
			O	4,5	1,5	0,92	0,4	9	4,1	1,5	0,78	0,7	17
CRatio	1 mesure	ÉAJT	N	1,3	0,3	0,62	0,2	14	1,3	0,3	0,47	0,2	17
			O	1,3	0,3	0,62	0,2	15	1,2	0,3	0,68	0,1	12
		ÉAJF	N	1,5	0,4	0,64	0,2	14	1,4	0,4	0,69	0,2	14
			O	1,5	0,4	0,84	0,1	10	1,4	0,4	0,73	0,2	14
	3 mesures	ÉAJT	N	1,3	0,3	0,61	0,2	13	1,3	0,3	0,46	0,2	16
			O	1,3	0,3	0,61	0,2	15	1,2	0,3	0,73	0,1	11
		ÉAJF	N	1,4	0,3	0,73	0,2	12	1,4	0,3	0,70	0,2	13
			O	1,4	0,4	0,83	0,1	10	1,4	0,4	0,75	0,2	13

PS : contrôle de la position de la sonde (N : non; O : oui); M : moyenne; ET : écart-type; ID : Indice de dépendabilité; ESM : erreur standard de mesure ; ESM (%) : ESM en pourcentage de la moyenne (M), Les ID $\geq 0,75$ et ESM (%) $\leq 10\%$ (critères d'excellente fidélité) sont identifiés en caractères gras.