

**Dérivation de règles de prédiction clinique
pour dépister les patients
ayant une lombalgie non aiguë
manifestant un succès thérapeutique
lors d'un programme d'exercice
de stabilisation lombaire**

Christian Larivière
Marie-France Coutu
Sharon M. Henry
Richard Preuss
Nicolas Roy
Khalil Rabhi

RAPPORTS
SCIENTIFIQUES

R-1151-fr
Version corrigée

NOS RECHERCHES travaillent pour vous !

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.

Mission

Dans l'esprit de la Loi sur la santé et la sécurité du travail (LSST) et de la Loi sur les accidents du travail et les maladies professionnelles (LATMP), la mission de l'IRSST est de :

Contribuer à la santé et à la sécurité des travailleuses et travailleurs par la recherche, l'expertise de ses laboratoires, ainsi que la diffusion et le transfert des connaissances, et ce, dans une perspective de prévention et de retour durables au 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.irstst.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 CNESST (preventionautravail.com)
- au bulletin électronique InfoIRSST

Dépôt légal

Bibliothèque et Archives nationales du Québec, 2022
ISBN 978-2-89797-253-0 (PDF) Version corrigée
ISBN 978-2-89797-219-6 (PDF) Version originale

© Institut de recherche Robert-Sauvé en santé et en sécurité du travail, 2022

IRSST - Direction des communications, de la veille et de la mobilisation des connaissances
505, boul. De Maisonneuve Ouest
Montréal (Québec)
H3A 3C2
Téléphone : 514 288-1551
publications@irstst.qc.ca
www.irstst.qc.ca

Dérivation de règles de prédiction clinique pour dépister les patients ayant une lombalgie non aiguë manifestant un succès thérapeutique lors d'un programme d'exercice de stabilisation lombaire

Christian Larivière^{1,7}, Marie-France Coutu², Sharon M. Henry³, Richard Preuss^{4,7}, Nicolas Roy^{5,7} et Khalil Rabhi⁶

¹ IRSST

² Université de Sherbrooke

³ Université du Vermont, Burlington, États-Unis

⁴ Université McGill

⁵ Chargé de cours, Université de Sherbrooke; Physiothérapeute clinicien en pratique privée, Physio Extra

⁶ Statisticien consultant indépendant

⁷ Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain (CRIR)

RAPPORTS
SCIENTIFIQUES

R-1151-fr

Version corrigée

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.

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

Erratum

Page 46, figure 6 : Le test TPP-portée a été remplacé par le TPP-flexions.

Page 47, tableau 13 : Le titre du tableau se lit maintenant comme suit : Nombre de participants des groupes succès et échec en fonction du nombre de tests positifs de la RPC de succès à T34 (modèle 2 avec 4 variables de classe A).

Page 47, tableau 13 : L'entrée Nombre de tests positifs est passé à 4.

Page 47, tableau 14 : Le titre du tableau se lit maintenant comme suit : Indicateurs de performance diagnostique (intervalles de confiance à 95 %) associés au nombre de prédicteurs présents dans la RPC à T34 (modèle 2 à 4 variables de classe A).

Page 47, tableau 14 : L'entrée Nombre de prédicteurs présents est passé à 4.

Page 62 : La 1^{re} phrase du dernier paragraphe se lit maintenant comme suit : Le test de performance physique TPP-portée a été sélectionné dans la RPC finale à T8, ceci avec un seuil $\geq 0,75/1$.

Page 67 : La 1^{re} phrase du 2^e paragraphe se lit maintenant comme suit : Le temps pris pour exécuter cinq flexions/extensions du tronc (variable TPP-flexions $\geq 9,91$ s) a été retenu dans la RPC finale à T34.



ÉVALUATION PAR DES PAIRS

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.

REMERCIEMENTS

Nous désirons remercier le Centre de recherche interdisciplinaire en réadaptation du Montréal métropolitain (CRIR) – pavillon de l'Institut de réadaptation Gingras-Lindsay de Montréal (IRGLM) – de l'Institut universitaire sur la réadaptation en déficience physique de Montréal (IURDPM), qui est au sein du Centre intégré universitaire de santé et de services sociaux du Centre-Sud-de-l'Île-de-Montréal (CCSMTL), pour son soutien en matière d'infrastructure de recherche.

Ce projet n'aurait pas vu le jour sans la participation de huit cliniques privées PhysioExtra de la région de Montréal : (1) Montréal-Fleury (2171, rue Fleury Est, Montréal) ; (2) Laval (3, boulevard Samson Sainte-Dorothée, Laval) ; (3) Gym du Plateau Mont-Royal (1688, avenue Mont-Royal, Montréal) ; (4) Regina-Assumpta (1750, rue Sauriol Est, Montréal) ; (5) Villeray (8560, rue Saint-Hubert, bureau 210, Montréal) ; (6) Gym Saint-Henri (4330, rue Saint-Jacques, Montréal) ; (7) Rosemont (2957 rue Bélanger, Montréal) ; (8) Cartierville (5777, boul. Gouin Ouest, bureau 230, Montréal). Nous exprimons notre gratitude à l'endroit des personnes qui nous ont ouvert les portes de ces cliniques, soit Pascal Gagnon (directeur général de PhysioExtra) et Magali Frève-Marchand (chef de communauté de pratique en physiothérapie), de même que les 11 physiothérapeutes qui ont été formés pour prodiguer les soins aux patients (Florian Blache, Camille Dubois-Chalifoux, Antoine Léger, Karine Malo, Catherine Mireault-Germain, Laetitia Mongeau, Sanand Patel, Élyse St-Pierre-Béluse, Vinh Tang, Ariane Viau, Maximilian Zucchi). La contribution du personnel de PhysioExtra à l'avancement de la science est tout à leur honneur.

Des remerciements tout particuliers s'adressent à l'équipe d'assistants de recherche pour la cueillette des données au Laboratoire d'évaluation des fonctions neuromusculaires du tronc de l'IURDPM pour la qualité de leur travail, soit Cynthia Appleby, Mélanie Clavet et Amir Chitour. Leur rigueur a permis non seulement une perte minimale, voire négligeable de mesures, mais aussi d'assurer une très bonne qualité de ces mesures.

Nous voulons aussi remercier Hakim Mecheri de l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) pour l'analyse des très nombreux signaux recueillis lors des séances en laboratoire, cela à plusieurs reprises pour satisfaire les nombreuses demandes des chercheurs et des étudiants gradués engagés dans ce projet. Les analyses statistiques très spécifiques de Djamal Berbiche (statisticien senior du Département des sciences de la santé communautaire, Faculté de médecine et des sciences de la santé, Université de Sherbrooke) ont également été essentielles à la compréhension de l'adhésion aux exercices faits à domicile.

Finalement, nous voudrions souligner le soutien technique qui a été offert par Daniel Marineau et Michel Goyette, ainsi que le suivi administratif du projet qui a été assuré par Lucille Benoît.

TABLE DES MATIÈRES

SOMMAIRE	VI
ACRONYMES, SIGLES ET ABRÉVIATIONS.....	IX
INTRODUCTION	1
Problématique générale des lombalgies en réadaptation	1
Objet de l'étude.....	2
1. ÉTAT DES CONNAISSANCES	3
1.1 Concept de stabilité lombaire	3
1.2 Quels patients répondent le mieux aux exercices de stabilisation lombaire ?.....	4
1.3 Quel type de programme d'exercice de stabilisation lombaire devrait être utilisé ?.....	5
1.4 Quels mécanismes d'action entrent en jeu ?	9
1.5 Résultats de l'étude préliminaire	13
2. OBJECTIFS DE RECHERCHE	16
3. MÉTHODOLOGIE	18
3.1 Devis expérimental	18
3.2 Participants	19
3.3 Intervention.....	22
3.4 Évaluations	23
3.4.1 Mesures de résultats (OUT)	23
3.4.2 Mesures cliniques de physiothérapie (PHT)	23
3.4.3 Mesures psychologiques (PSY).....	24
3.4.4 Mesures neuromusculaires (NRM)	26
3.5 Analyses statistiques	28
3.5.1 Objectif 1 – Dérivation de la RPC de succès (hypothèse 1).....	28
3.5.2 Objectif 2 – Étude des mécanismes d'action (hypothèses 2 et 3)	31
4. RÉSULTATS	33
4.1 Recrutement des participants	33
4.2 Effets du programme d'exercice.....	34
4.3 Objectif 1 – Dérivation de RPC de succès	36

4.3.1	Sélection des variables prédictives potentielles	36
4.3.2	Analyses multivariées	36
4.4	Objectif 2 – Étude des mécanismes d'action.....	47
4.4.1	Évaluation physique (mesures PHT)	47
4.4.2	Évaluation psychologique (mesures PSY)	52
4.4.3	Évaluation neuromusculaire (mesures NRM).....	57
5.	DISCUSSION	60
5.1	Objectif 1 – Dérivation des RPC de succès.....	60
5.1.1	Modèles ne comportant que des variables de classe A	61
5.1.2	Modèles comportant des variables de classe A et B.....	67
5.1.3	Modèles comportant des variables de classe A, B et C	70
5.1.4	Quelle RPC choisir ? Pour prédire le succès à la fin du traitement (T8) ou au suivi de six mois après la fin du traitement (T34) ?.....	71
5.2	Objectif 2 – Étude des mécanismes d'action.....	72
5.2.1	Évaluation physique (mesures PHT)	72
5.2.2	Évaluation psychologique (mesures PSY)	73
5.2.3	Évaluation neuromusculaire (mesures NRM).....	75
5.3	Forces et limites.....	77
	CONCLUSION.....	79
	BIBLIOGRAPHIE	80
ANNEXE A.	Description sommaire des mesures	101
ANNEXE B.	Résultats (objectif 1 : dérivation des RPC).....	108

LISTE DES TABLEAUX

Tableau 1.	Première méthode d'estimation de la taille d'échantillon	20
Tableau 2.	Deuxième méthode d'estimation de la taille d'échantillon.....	20
Tableau 3.	Description des caractéristiques démographiques, anthropométriques et cliniques des 110 patients	34
Tableau 4.	Effet du temps/traitement sur les mesures de résultats primaires (OUT), chez les 100 patients qui ont fait le suivi de 6 mois	35
Tableau 5.	Mesures de résultats primaires (OUT) chez les 110 patients qui ont complété le programme d'exercice de huit semaines, classés selon le niveau de succès du traitement.....	35
Tableau 6.	Mesures de résultats primaires (OUT) chez les 100 patients qui ont complété les mesures prises à six mois de suivi (T34), classés selon le niveau de succès du traitement.....	36
Tableau 7.	Les coefficients des modèles prédictifs de succès à T8 et leurs intervalles de confiance à 95% (IC 95%).....	38
Tableau 8.	Indicateurs de performance diagnostique et d'ajustement des modèles prédictifs à T8 (IC 95 %).	39
Tableau 9.	Nombre de participants des groupes succès et échec en fonction du nombre de tests positifs de la RPC de succès à T8 (modèle 3 avec 4 variables de classe A).....	40
Tableau 10.	Indicateurs de performance diagnostique (intervalles de confiance à 95 %) associés au nombre de prédicteurs présents dans la RPC de succès à T8 (modèle 3 à 4 variables de classe A)	40
Tableau 11.	Les coefficients des modèles prédictifs de succès à T34 et leurs intervalles de confiance à 95 % (IC 95 %).....	44
Tableau 12.	Indicateurs de performance diagnostique et d'ajustement des modèles prédictifs à T34 (IC 95 %).	45
Tableau 13.	Nombre de participants des groupes succès et échec en fonction du nombre de tests positifs de la RPC de succès à T34 (modèle 2 avec 4 variables de classe A).	47
Tableau 14.	Indicateurs de performance diagnostique (intervalles de confiance à 95 %) associés au nombre de prédicteurs présents dans la RPC à T34 (modèle 2 à 4 variables de classe A)	47
Tableau 15.	Effet du sous-groupe et du temps/traitement sur les mesures PHT (variables continues) recueillies aux temps T0 et T8.....	49

Tableau 16.	Effet du temps/traitement sur les mesures PHT (variables dichotomiques).....	51
Tableau 17.	Effet de sous-groupe (SG) et du temps/traitement (T) sur les mesures PSY (variables continues) recueillies aux temps T0, T4, T8 et T34.....	53
Tableau 18.	Effet du sous-groupe (SG) et du temps/traitement (T) sur les mesures PSY (variables continues) recueillies aux temps T1, T4 et T8.....	54
Tableau 19.	Modèle multivarié de l'adhésion au programme d'exercice à T8 et T34	56
Tableau 20.	Modèle multivarié du changement global (ChangGlob) perçu à T8 et T34	57
Tableau 21.	Description des caractéristiques démographiques et anthropométriques des sujets sains et lombalgiques qui ont réalisé les tests neuromusculaires	59
Tableau 22.	Description des mesures de résultats (OUT) des sujets lombalgiques qui ont réalisé les tests neuromusculaires	59
Tableau 23.	Comparaison entre les performances des différentes RPC de succès développées pour différents traitements offerts en physiothérapie.	62

LISTE DES FIGURES

Figure 1.	Le modèle peur/évitement.....	11
Figure 2.	Le modèle de stabilisation lombaire.	12
Figure 3.	Quatre catégories de mesures (OUT, PHT, PSY, NRM) ont été recueillies chez les patients à différents moments (T0, T4, T8) et après (T34) le programme d'exercice. Les mesures PHT et NRM sont plus coûteuses et n'ont donc été collectées qu'à T0 et T8.	19
Figure 4.	Diagramme de suivi des participants, de la phase de recrutement jusqu'à la fin de l'étude.	33
Figure 5.	Tests PHT retenus dans la RPC de succès à la fin du traitement (au temps T8).	41
Figure 6.	Tests PHT retenus dans la RPC de succès pour le suivi de six mois (au temps T34).....	46
Figure 7.	Illustration des interactions SOUS-GROUPE × TEMPS obtenues pour les variables PHT de type continu.	50
Figure 8.	Illustration des interactions GROUPE × TEMPS obtenues pour les variables PSY de type continu.	55

SOMMAIRE

Les exercices physiques permettent de réduire la douleur et les incapacités chez les gens souffrant d'une lombalgie non aiguë (> 3-4 semaines), mais ces effets sont limités. Pour rehausser l'efficacité de ce type d'intervention, il faut déterminer quels patients répondent le mieux à chaque modalité d'exercice (arrimage patient/intervention) et pour quelles raisons (mécanismes sous-jacents). Dans la présente recherche, l'intérêt est porté à un programme d'exercices de stabilisation lombaire (PESL), une modalité active d'exercice gagnant en crédibilité et en popularité. Deux objectifs ont été poursuivis : (1) finaliser l'étape de dérivation de règles de prédiction clinique (RPC) de succès¹ pour dépister, lors de l'examen clinique, les patients qui répondront bien à ces exercices ; (2) étudier les mécanismes (d'origines physique, psychologique et neuromusculaire) mis en action par ces exercices à l'aide de mesures plus spécifiques permettant de décrire les effets du traitement. Une étude préliminaire menée auprès de 48 patients avait produit les résultats suffisants pour poursuivre le recrutement de patients additionnels et ainsi atteindre les effectifs nécessaires pour répondre à ces objectifs.

À partir des résultats de l'étude préliminaire, 64 patients additionnels (douleur depuis plus de 4 semaines) devaient être recrutés pour atteindre un échantillon de 107 patients, soit la taille d'échantillon totale nécessaire sur le plan statistique. Le programme d'exercice était réalisé sur 8 semaines (2 séances/semaine) dans des cliniques de physiothérapie. Les principales mesures de résultats [douleur ; perceptions d'incapacité (échelle Oswestry)], de même que plusieurs mesures par questionnaire (mesures psychologiques [PSY] associées à la douleur et à l'adhésion au traitement) ont été recueillies au début (T0), aux semaines 4 (T4), 8 (T8 - fin) du programme d'exercice ainsi qu'à six mois post-traitement (T34). Les autres mesures réalisables en milieu clinique et donc susceptibles d'être retenues pour le développement des RPC (objectif 1), soient les tests physiques (mesures de l'examen clinique en physiothérapie [PHT]) faits à l'aide d'un examen clinique, ont été réalisées à T0 et T8. Ces tests physiques comprenaient des tests d'instabilité articulaire (n = 4), de flexibilité (n = 6), de déficiences du contrôle sensorimoteur (n = 8), de performance physique (n = 4) et d'endurance musculaire (n = 3). Six tests en laboratoire ont aussi été réalisés à T0 et T8 pour étudier les mécanismes d'action d'origine neuromusculaire (mesures neuromusculaires [NRM] ; pour l'objectif 2), chez un sous-échantillon de 77 patients.

¹ Une RPC de succès et une RPC d'échec étaient visés au départ. Au cours de la présente étude, il a été jugé plus utile de se concentrer que sur la RPC de succès, mais au lieu de ne considérer que le succès immédiat à la fin du programme clinique, une RPC de succès a également été dérivée pour la prédiction de succès lors du suivi à six mois.

Pour les deux objectifs, les analyses consistaient d'abord à former trois sous-groupes de patients en fonction de leur niveau de succès suivant le PESL (1. succès ; 2. amélioration cliniquement significative ; 3. échec), en utilisant le questionnaire Oswestry, mesurant la perception d'incapacité liée à un mal de dos. Ces analyses consistaient ensuite à comparer les deux groupes extrêmes (sous-groupes succès (GS) vs échec (GE)) afin de s'assurer qu'ils sont bien distinctifs. D'ailleurs, autant à T8 qu'à T34, les tailles d'effet du PESL correspondant au GS étaient beaucoup plus élevées que pour le GE. Ceci était attendu pour la mesure de perception d'incapacité, mais aussi observé, bien qu'à un degré moindre, pour l'intensité de la douleur.

Résultats et discussion en lien avec l'objectif 1 (développement des RPC) : les deux RPC de succès (à T8 et T34), toutes deux composées de tests ayant un lien théorique direct avec le concept d'instabilité lombaire, ont démontré une très bonne valeur prédictive. La RPC de succès à T8, dérivée avec 99 patients (GS + GE) pour prédire le succès à la fin du traitement clinique de 8 semaines a retenu quatre variables de l'examen physique. Lorsqu'au moins deux de ces variables sont positives, le rapport de vraisemblance positif est très élevé ($LR+ = 17,9$), permettant de rehausser la probabilité de succès du PESL de 49 % (sans utilisation de la RPC) à 96 % à T8, soit une amélioration de 47 % de cette probabilité de succès. Ces indicateurs de performance diagnostique sont supérieurs à la RPC préliminaire dérivée par Hicks *et al.*, (2005) pour un PESL, composée de quatre variables, et offrant un $LR+$ de 4,0 et une probabilité de succès passant de 33 (sans RPC) à 67 % à T8 (amélioration de 34 %).

La RPC de succès à T34, dont la dérivation clinique est basée sur un échantillon de 89 patients pour prédire le succès après une période de suivi de six mois, a retenu trois variables prédictives ajustées sur la variable âge. Lorsqu'au moins trois de ces variables sont positives, le rapport de vraisemblance positif est aussi très élevé ($LR+ = 17,0$), permettant de rehausser la probabilité de succès du PESL de 53 % (sans utilisation de la RPC) à 85 % à T34, soit une amélioration de 32 % de cette probabilité de succès. En comparaison avec la RPC préliminaire de Hicks *et al.* (2005), dérivée pour prédire le succès à T8 (à la fin du programme clinique), la RPC à T34 offre un $LR+$ beaucoup plus élevé (17,0 vs 4,0), mais un changement équivalent en termes de prédiction du succès (amélioration de 32 % vs 34 %). Il est important d'apprécier qu'il s'agisse de la première RPC dans le domaine de la physiothérapie, quant à la prédiction du succès après une période de suivi, ici s'échelonnant sur six mois.

Résultats et discussion en lien avec l'objectif 2 (étude des mécanismes d'action) : Les comparaisons entre les GS et GE dans le temps ont révélé que plusieurs mécanismes physiques (mesures de l'examen clinique) et psychologiques (mesures par questionnaires) ont le potentiel d'expliquer le succès au PESL, ce qui n'était pas le cas pour les mécanismes de nature neuromusculaire (mesures en laboratoire). Ainsi, le PESL permettrait d'améliorer de façon plus marquée plusieurs indicateurs cliniques (p. ex. tests physiques d'endurance musculaire) et psychologiques (p. ex. peurs et croyances face à

l'activité physique) chez le GS que chez le GE. Ceci supporte l'idée que le PESL permet une exposition graduelle à l'activité qui est favorable à la réduction des craintes liées à la douleur et au mouvement, comme le propose le modèle d'évitement de la peur.

Finalement, l'étude plus spécifique de l'adhésion au programme d'exercices à domicile suggère que certains des mécanismes de nature psychologique en question joueraient un rôle dans cette adhésion, ce qui a son tour aurait un impact sur le succès du PESL.

En conclusion, la présente recherche a permis de compléter avec succès l'étape de dérivation de deux RPC de succès d'un PESL, soit pour prédire le succès à la fin d'un traitement de huit semaines (RPC à T8) ou bien après un suivi de six mois (RPC à T34). Ceci justifie la poursuite du développement de ce type d'outil permettant un meilleur arrimage entre les patients et ce programme d'exercices en particulier. De plus, l'étude des mécanismes d'action suggère que ce programme d'exercice engage des mécanismes d'action bénéfiques pour les patients, ceci autant sur le plan physique que psychologique.

ACRONYMES, SIGLES ET ABRÉVIATIONS

Acronyme	Définition
ANOVA	Analyse de variance
APA	Ajustements posturaux anticipatoires
C7	7 ^e vertèbre cervicale
CCIC	Coefficient de corrélation intraclasse
ChangGlob	Perception du changement global de l'état de santé (lombalgie)
ECR	Essai clinique randomisé
IMC	Indice de masse corporelle
Int-Douleur	Intensité de la douleur
L1, L3, L4, L5	1 ^{re} , 3 ^e , 4 ^e et 5 ^e vertèbres lombaires
LR- et LR+	Rapports de vraisemblance (<i>likelihood ratio</i>) négatif et positif
LuM	Multifides lombaires (<i>Lumbar multifidus</i>)
NRM	Mesures neuromusculaires
OI	Oblique interne
OE	Oblique externe
OR	Rapport de cote (<i>Odd ratio</i>)
OSW	Perception d'incapacité – Questionnaire Oswestry
OUT	Mesures de résultat (<i>outcomes</i>)
PESL	Programmes d'exercices de stabilisation lombaire
PHT	Mesures de l'examen clinique en physiothérapie
PSY	Mesures psychologiques
ROC	<i>Receiver operating characteristic</i>
RPC	Règle de prédiction clinique
Satisfaction-PT	Satisfaction du patient concernant les services de physiothérapie
Se	Sensibilité
Sp	Spécificité
T0, T1, T4, T8 et T34	Temps de mesure aux semaines 0, 1, 4, 8 et 34
T12	12 ^e vertèbre thoracique
TrA	Transverse de l'abdomen
VP- et VP+	Valeurs prédictives négative et positive

* La liste détaillée des abréviations correspondant aux variables PHT, PSY et NRM est présentée à l'annexe A.

INTRODUCTION

Problématique générale des lombalgies en réadaptation

Les lombalgies sont courantes, avec une prévalence au cours de la vie variant de 49 à 70 %. Parmi 291 problèmes de santé, elle est classée au premier rang sur le plan de l'incapacité et au sixième en matière d'impact économique (Hoy *et al.*, 2014). La lombalgie chronique, qui représente une petite proportion de l'ensemble des lombalgies, mais qui est à l'origine de la plupart des coûts (Hashemi *et al.*, 1997), a vu sa prévalence passer de 3,9 % en 1992 à 10,2 % en 2006 (Freburger, Holmes, *et al.*, 2009). Au Québec, les statistiques de la Commission de la santé et de la sécurité du travail (CSST) montrent que les lésions de la colonne vertébrale représentent 30 % de l'ensemble des lésions professionnelles indemnisées entre 2008 et 2011, la région lombaire étant impliquée dans plus de 60 % de ces cas (CSST, 2012). Les individus souffrant de lombalgie non spécifique, diagnostic basé sur l'exclusion d'une pathologie spécifique, représentent environ 90 % des personnes souffrant de lombalgie (Koes *et al.*, 2006). Bien entendu, un « diagnostic » non spécifique entraîne un traitement non spécifique (conservateur).

Des revues systématiques indiquent que les exercices physiques sont efficaces pour réduire la douleur et l'incapacité physique chez les personnes souffrant de lombalgie « non aiguë » (Searle *et al.*, 2015 ; Smidt *et al.*, 2005), ainsi que pour prévenir la lombalgie (Bigos *et al.*, 2009) ; ce dernier point est encourageant étant donné le taux élevé de récurrence. Il n'est donc pas étonnant que les exercices physiques figurent parmi les traitements conservateurs les plus recommandés pour les personnes souffrant d'une lombalgie non spécifique en période subaiguë (d'une durée supérieure à 4 semaines) (Rossignol, 2006) ou chronique (de plus de 12 semaines) (Rossignol, 2006 ; Foster *et al.*, 2018). D'ailleurs, il est prescrit à environ 50 % de tous les patients qui consultent un médecin, un physiothérapeute ou un chiropraticien (Freburger, Carey, *et al.*, 2009). Cependant, les recherches peinent à démontrer la supériorité d'un type de programme d'exercice par rapport à un autre (Hayden *et al.*, 2005 ; Smith, B. E. *et al.*, 2014). De plus, les effets du traitement pour l'exercice, dans le cas de la lombalgie non aiguë et non spécifique, sont au mieux modérés. Identifier quel programme d'exercice est le plus efficace semble ne plus être la question de recherche la plus pertinente à se poser.

Les patients atteints d'une lombalgie non aiguë et non spécifique sont loin de constituer un groupe homogène en termes de facteurs physiologiques et psychologiques. D'ailleurs, l'idée selon laquelle des patients ayant différentes pathologies ou troubles puissent répondre inégalement à différentes modalités d'exercice (Leboeuf-Yde et Manniche, 2001) a désormais des fondements scientifiques. En effet, quelques essais cliniques randomisés (ECR) ont démontré qu'un meilleur appariement entre certains sous-groupes de patients et certaines modalités d'intervention favorise de meilleurs résultats cliniques (Brennan *et al.*, 2006 ; Fritz *et al.*, 2003 ; Long *et al.*, Fung, 2004). Ainsi, le développement

de méthodes de classification des patients lombalgiques, comme ceux susceptibles de répondre à des interventions spécifiques d'exercice physique, a été identifié comme une priorité de recherche par plusieurs groupes d'experts (Airaksinen *et al.*, 2004; American Physical Therapy Association [APTA], 2000; Costa *et al.*, 2013; Savigny, Watson *et al.*, 2009), ainsi que la priorité n° 1 des intervenants de première ligne (chiropraticiens, médecins, physiothérapeutes) qui traitent les patients souffrant d'une lombalgie (Henschke *et al.*, 2007).

Objet de l'étude

Au-delà de la simple comparaison de programmes d'exercice à l'aide d'essais cliniques randomisés (ECR), il importe de déterminer : (1) les patients qui répondent le mieux à un programme d'exercice donné (arrimage patient/intervention) ; (2) les mécanismes sous-jacents qui contribuent à la réussite du traitement.

Une modalité active d'exercice, soient les programmes d'exercices de stabilisation lombaire (PESL), dispose de solides assises scientifiques, et ce depuis plusieurs années (Barr *et al.*, 2007 ; McGill, 2002 ; Richardson *et al.*, 2004). D'ailleurs, une méta-analyse soutient l'efficacité des deux composantes qui constituent les PESL, soient les exercices de force/résistance et les exercices de coordination/stabilisation (Searle *et al.*, 2015). Par contre, comme le souligne la plus récente revue systématique Cochrane sur le sujet (Saragiotto *et al.*, 2016), un arrimage patient/intervention fondé sur des preuves n'est toujours pas établi pour ce type de PESL.

Les PESL adhèrent aux principes psychologiques de l'exposition graduelle à l'activité physique, lesquels permettent une réduction de la peur associée à la douleur et au mouvement (Vlaeyen et Linton, 2000). Il a été démontré qu'ils favorisent les changements neuroplastiques dans le cortex moteur, ainsi que les changements associés dans la coordination de certains muscles profonds tels que les transverses de l'abdomen (Tsao *et al.*, 2010). Il est également compatible avec les stratégies qui maximisent les résultats cliniques, car le programme d'exercice est actif, supervisé et implique un dosage élevé (Hayden *et al.*, 2005 ; Vuori, 2001). Bien qu'un dosage élevé ne puisse être atteint qu'avec des exercices supplémentaires effectués à domicile, cela est réalisable, car il ne nécessite pas d'équipement spécialisé et cela est conforme à l'objectif de développer l'efficacité personnelle du patient à réaliser les exercices de façon autonome. La popularité des exercices de stabilisation lombaire au sein de la communauté de physiothérapeutes souligne également la nécessité de cette recherche, considérant que plus de 94 % des travailleurs souffrant d'une lombalgie non aiguë sont référés pour la physiothérapie en première ligne (Lebeau, 2021). Dans la présente étude, il est important de préciser que les participants n'étaient pas nécessairement indemnisés par la Commission des normes, de l'équité, de la santé et de la sécurité du travail (CNESST). La population à l'étude devait essentiellement avoir une lombalgie non aiguë et satisfaire les critères d'inclusion et d'exclusion détaillés à la section 4.2 (Participants)

1. ÉTAT DES CONNAISSANCES

1.1 Concept de stabilité lombaire

« La stabilité est un terme très chargé dont la définition est instable » (Reeves *et al.*, 2019) et de plus, dont les mesures sont imparfaites. En effet, la stabilité lombaire est un concept ambigu qui s'est mérité plusieurs interprétations (Reeves *et al.*, 2007 ; Eisenstein, 1999) en fonction des perspectives considérées (mécanique, clinique, contrôle sensorimoteur). Cependant, bien que des efforts soient faits pour mieux la décrire avec des connaissances interdisciplinaires (Reeves *et al.*, 2019), aucune définition ne semble faire consensus. Voici donc quelques définitions permettant d'identifier les éléments clés de ce concept, comme détaillé et illustré par Reeves *et al.* (2011) :

1. La stabilité mécanique est présente lorsqu'un système revient à son état initial non perturbé (position statique ou trajectoire de mouvement prévue) à la suite d'une perturbation mécanique. La performance du système décrit la rapidité avec laquelle le système revient à son état initial non perturbé. La robustesse du système décrit la gamme de perturbations pour lesquelles le système est stable.
2. D'un point de vue clinique, cependant, il ne suffit pas qu'une articulation revienne simplement à son état initial non perturbé. Cela doit se faire sans causer de blessure à l'articulation ou aux tissus périarticulaires. Le concept de stabilité clinique englobe donc la stabilité mécanique, la performance (le système neuro-musculosquelettique doit être capable de stabiliser l'articulation assez rapidement pour prévenir les blessures) et la robustesse (le système neuro-musculosquelettique doit être capable de répondre à toutes les perturbations que l'articulation pourrait rencontrer pendant l'activité). Cela peut être compromis par :
 - des blessures ou anomalies antérieures de l'articulation (p. ex., blessure ligamentaire ou congruence congénitale de l'articulation) ;
 - un fonctionnement sous-optimal de la musculature (force, endurance, puissance, flexibilité et rigidité) ;
 - un fonctionnement sous-optimal du système nerveux, qui peut compromettre :
 - le contrôle sensorimoteur avec commandes prédictives (ajustements posturaux anticipatoires), nécessaire pour que l'articulation reste stable pendant les perturbations mineures, sans avoir besoin de changements entraînés par rétroaction dans l'activation musculaire,
 - une commande motrice entraînée par rétroaction (réponses réflexes) ; les réponses retardées ou imprécises basées sur la rétroaction peuvent permettre au système de devenir instable avant qu'une réponse appropriée puisse être générée (mauvaise performance) ou causer une réponse inappropriée conduisant à d'autres perturbations ou instabilités (faible robustesse).

Donc, une instabilité clinique existe quand le système neuro-musculosquelettique ne peut pas empêcher (ou cause par inadvertance) des mouvements physiologiques ou accessoires indésirables à l'articulation, pendant des activités et à des niveaux de force qui devraient normalement être tolérés, menant à la blessure de tissus et/ou à la douleur nociceptive.

1.2 Quels patients répondent le mieux aux exercices de stabilisation lombaire ?

De manière à améliorer l'efficacité des PESL, des règles de prédiction clinique (RPC) de succès et d'échec doivent être développées pour prescrire ce programme d'exercice aux patients appropriés (Fritz, 2009 ; Stanton *et al.*, 2010). Il est important de savoir qu'une RPC ne peut être diffusée et utilisée qu'après la réalisation d'études de validation et d'analyse d'impact (Childs et Cleland, 2006 ; deux étapes que nous prévoyons réaliser ultérieurement si cette première étape de développement (stade de dérivation) mène à des résultats positifs. Des travaux préliminaires de dérivation de deux RPC (succès et échec) ont déjà été faits pour un PESL (Hicks *et al.*, 2005). Leur validité conceptuelle a été testée avec succès (Hebert *et al.*, 2010 ; Teyhen *et al.*, 2007), mais leur validation plus formelle a échoué en raison d'une puissance statistique insuffisante (Rabin *et al.*, 2014). Ces résultats montrent qu'il est possible de dériver et, du moins partiellement, valider une RPC, mais l'étape de dérivation (Hicks *et al.*, 2005) souffre de limites importantes :

- Étant une étude préliminaire (titre : Preliminary development of a clinical prediction rule (Hicks *et al.*, 2005), 54 patients avaient été recrutés, ce qui est insuffisant sur la base de l'analyse de puissance, comme nous l'avons estimée dans la section méthodologie, soit le double (n = 107 patients).
- L'échantillon de patients était hétérogène en matière de durée des symptômes (phases aiguë, subaiguë et chronique), ce qui n'est pas en phase avec le fait que les exercices physiques ne sont pas efficaces au cours de la phase aiguë de la lombalgie (van Tulder *et al.*, 2000) et que la phase subaiguë est plus propice pour déterminer les patients à risque d'incapacité (Waddell *et al.*, 2003).
- Bien que « les mouvements aberrants », observés lors de la flexion avant du tronc, soient une mesure composite considérant cinq signes cliniques portant sur le contrôle dysfonctionnel du mouvement, d'autres tests évaluant le contrôle actif des mouvements de la région lombo-pelvienne, basés sur les théories de Sahrman (2002) et de O'Sullivan (2005), peuvent être considérés.
- Une seule variable psychologique (peur/évitement) a été prise en considération. Une classification basée sur un modèle de prévention de l'incapacité est recommandée afin de clairement reconnaître non seulement l'hétérogénéité de cette population en fonction de leur physique, mais aussi sur le plan de leurs caractéristiques psychologiques (McCarthy *et al.* 2004 ; O'Sullivan, 2005). Le modèle psychologique peur/évitement (figure 1) est l'un des modèles les plus robustes pour expliquer la douleur à long terme et l'incapacité (Vlaeyen et Linton, 2000). En dehors de la peur/évitement en elle-même, d'autres variables psychologiques sont incluses dans ce modèle, telles que la dramatisation de la douleur et la détresse psychologique ou les symptômes dépressifs qui sont

susceptibles d'affecter le succès de tout programme d'exercice. En outre, cela pourrait être fait en vertu de leur effet sur l'adhésion au programme d'exercice fait à domicile. Une faible adhésion est associée à de plus faibles résultats cliniques chez les patients lombalgiques (Hicks *et al.*, 2012 ; Iversen *et al.*, 2003 ; Liddle *et al.*, 2004), ce qui est aussi vrai pour les exercices de stabilisation lombaire (Mannion *et al.*, 2009 ; Nava-Bringas *et al.*, 2016). Plusieurs variables ont le potentiel de prédire l'adhésion aux exercices (voir méthodologie) et doivent être considérées comme des candidates potentielles de la RPC. D'ailleurs, l'adhésion étant un champ de recherche en soi, l'utilisation d'un cadre théorique est à préconiser.

- Le programme d'exercices mis en œuvre correspondait davantage à un programme de renforcement [école de pensée de McGill (2002)] qu'à un programme d'exercices où l'accent est mis uniquement sur le contrôle sensorimoteur [école de pensée australienne ; (Richardson *et al.*, 2004)]. Nous pensons qu'une approche hybride (l'approche australienne suivie par l'approche de McGill) est préférable, comme indiqué dans la section suivante.
- Finalement, une revue critique de la littérature portant sur les RPC développées pour les interventions en physiothérapie souligne que toutes ces RPC ne prédisent le succès qu'à la fin du traitement, alors qu'il serait très informatif de prédire ce succès après une période de suivi d'au moins six mois après le traitement (Beneciuk *et al.*, 2009).

1.3 Quel type de programme d'exercice de stabilisation lombaire devrait être utilisé ?

Il y a actuellement controverse en ce qui a trait au type de programme d'exercice qui permettrait de mieux restaurer la stabilité lombaire chez les patients dont la stabilité lombaire est compromise. Comme mentionné plus tôt, il y a deux grandes écoles de pensée : (1) l'approche australienne (Richardson *et al.*, 2004) soutenue par les principaux chercheurs australiens (J. Hides, P. Hodges, C. Richardson) et (2) l'approche de McGill (2002) défendue par un seul chercheur (S. McGill de U. de Waterloo, Canada).

L'approche australienne utilise les principes d'apprentissage sensorimoteur pour enseigner le contrôle et la coordination optimale de la colonne vertébrale. Il y a entraînement de la préactivation des muscles profonds du tronc, avec une progression vers des tâches statiques, dynamiques et fonctionnelles plus complexes qui intègrent l'activation des muscles profonds (ou locaux) et superficiels (ou globaux). L'idée principale est de parvenir à une bonne maîtrise des muscles profonds du tronc (transverses de l'abdomen [TrA], multifides lombaires [LuM]) pour contrôler la stabilité vertébrale avant de mettre à contribution les muscles superficiels du tronc. Pour ce faire, l'entraînement consiste à apprendre à recruter les TrA et LuM qui sont apparemment inhibés par la douleur et qui ne joueraient plus leur rôle de stabilisation. C'est que les TrA et LuM peuvent agir directement ou indirectement sur les vertèbres lombaires, ce qui est potentiellement favorable pour assurer la stabilité intersegmentaire (Barker *et al.*, 2006), contrairement aux muscles plus superficiels dont la vocation consiste davantage à

générer un mouvement tout en contribuant à la stabilité de la colonne lombaire dans son ensemble, c.-à-d. de façon moins spécifique. L'entraînement des TrA et des LuM se fait avec des exercices très fins et spécifiques (en contrôle sensorimoteur), permettant de faciliter leur contraction plus isolément, c.-à-d. avec une participation minimale des muscles plus superficiels. Les méthodes de rétroaction disponibles sont utilisées pour aider le patient à comprendre la nature de ces exercices très fins et spécifiques. Ces contractions isolées des TrA et LuM sont graduellement intégrées dans des exercices statiques faits au sol et, enfin, dans des tâches fonctionnelles plus dynamiques.

L'approche de McGill propose une forme plus simple d'entraînement en contrôle sensorimoteur dès le début du programme afin d'améliorer la conscience corporelle (proprioception). Les exercices statiques au sol sont introduits sans tarder, mais toujours graduellement. Ces exercices sont conçus pour : (1) rigidifier la région lombaire en contractant globalement (et non isolément), mais doucement la paroi abdominale tout en s'assurant de conserver une lordose lombaire neutre ; (2) surcharger les muscles et ainsi améliorer l'endurance des muscles du tronc ; (3) éviter d'exacerber la douleur en mettant trop de charges sur la colonne vertébrale. Ils consistent à soutenir différentes postures statiques de façon répétitive (programme basé sur des quotas), avec un niveau de difficulté croissant pour permettre une progression tout au long du programme. L'endurance des muscles du tronc est favorisée parce que la fatigabilité des muscles spinaux lombaires est un prédicteur d'un premier épisode de lombalgie (Biering-Sorensen, 1984) et d'incapacité à long terme liée à une lombalgie (Enthoven *et al.*, 2003). On suppose que la fatigue des muscles du dos augmenterait le risque d'erreurs neuromusculaires qui provoqueraient de brefs mouvements intervertébraux incontrôlés qui entraîneraient des lésions des tissus qui stabilisent ces vertèbres, comme le propose l'hypothèse de stabilité de la colonne vertébrale (Granata et Gottipati, 2008). En d'autres termes, l'amélioration de l'endurance des muscles du tronc contribuerait à améliorer les résultats cliniques et à protéger contre les récives, bien que ce dernier point n'ait pas encore été démontré.

Même si les chercheurs conviennent que tous les muscles du tronc (muscles profonds et superficiels de l'abdomen et du dos) jouent un rôle pour assurer la stabilité lombaire, une controverse persiste sur l'importance accordée à l'entraînement spécifique des muscles profonds par rapport aux muscles superficiels. Stuart McGill a fourni des preuves, sur la base d'études biomécaniques et sans prêter attention aux muscles profonds ou superficiels, qu'une contraction globale de la paroi abdominale fournit une plus grande stabilité lombaire qu'une contraction isolée du TrA, du moins sur le plan de la réduction du mouvement de la colonne vertébrale en réaction à des perturbations soudaines du tronc (Vera-Garcia *et al.*, 2007), ce qui a plus récemment été corroboré par l'étude de la rigidité lombaire (Larivière, Boucher, *et al.*, 2019). Cependant, ces preuves demeurent insuffisantes, car les mouvements intervertébraux ne peuvent pas être évalués à l'aide de ces techniques de mesure.

En ce qui a trait à l'approche australienne, il y a un nombre croissant de preuves neurophysiologiques suggérant que la musculature profonde de la colonne vertébrale est altérée chez les personnes souffrant de lombalgie ou qui ont eu un épisode récent (c.-à-d. sans douleur au moment du test). Différentes observations interreliées ont été étayées. Tout d'abord, une atrophie des LuM, qui est ipsilatérale aux symptômes (Barker *et al.* 2004 ; Hides *et al.* 1994) et spécifique à ce groupe de muscles par rapport aux muscles superficiels (Danneels *et al.*, 2000), a été observée. Deuxièmement, une réduction du contrôle volontaire des LuM et TRA a été documentée (Henry et Teyhen, 2007). Enfin, une activation retardée ou moindre des TrA a été observée chez les patients atteints d'une lombalgie comparativement à des sujets sains, en réponse à une perturbation (mouvement rapide du bras) qui met à l'épreuve la stabilité vertébrale. Ces résultats ont parfois été observés simultanément pour l'oblique interne (Hodges et Richardson, 1999), mais pas pour l'oblique externe qui est plus superficiel (Ferreira *et al.*, 2004 ; Hodges et Richardson, 1999). Des résultats similaires ont été récemment observés chez les personnes atteintes de lombalgie récurrente par rapport à des sujets sains pour les LuM (MacDonald *et al.*, 2009 ; MacDonald *et al.*, 2010). Ces activations musculaires préparatoires, appelées ajustements posturaux anticipatoires (APA), sont préprogrammées par le système nerveux central afin de minimiser les perturbations posturales (Tsao *et al.*, 2009) et par conséquent donnent un aperçu unique du contrôle neuronal central de la posture. Bien que ces déficiences (atrophie, moindre contrôle volontaire, APA retardés) ont aussi été observées pour les muscles superficiels du tronc, les LuM et TrA ont montré des résultats anormaux de façon plus consistante. Par contre, il y a aussi des études qui n'ont pas observé ces déficiences (Boucher *et al.*, 2018 ; Brooks *et al.*, 2012 ; Mehta *et al.*, 2017), confirmant que les personnes atteintes de lombalgie ne représentent pas une population hétérogène. D'ailleurs, le dysfonctionnement du tronc et des muscles abdominaux peut être influencé par des facteurs psychologiques (Masse-Alarie *et al.*, 2016b).

Les patients souffrant de lombalgie présentent également des modifications fonctionnelles (réorganisation) des propriétés neuronales du système sensorimoteur représentant les muscles les plus touchés par la douleur, telles qu'une réduction de l'activation motrice corticale des muscles lombaires (Strutton *et al.*, 2005), un déplacement de la représentation des muscles du bas du dos dans le cortex somatosensoriel (Flor *et al.* 1997) et un déplacement postérieur et latéral de la représentation ainsi qu'une plus grande représentation du TrA dans le cortex moteur primaire (Tsao *et al.*, 2008). Les APA du TrA sont en corrélation avec la mesure dans laquelle la zone de représentation de ce muscle dans le cortex moteur est remodelée chez les personnes souffrant de douleurs lombaires récurrentes, telle que mesurée par stimulation magnétique transcrânienne (Tsao and Hodges 2008; Tsao *et al.* 2010). De plus, l'activité cérébrocorticale est modifiée avant les mouvements du bras nécessitant un APA chez les personnes souffrant d'une lombalgie chronique (Jacobs *et al.* 2010). Des recherches plus récentes continuent de montrer un lien entre la lombalgie chronique, les changements corticomoteurs et le contrôle sensorimoteur des muscles du tronc (Masse-Alarie *et al.*, 2016a, 2017).

La prémisse de l'approche australienne est que les exercices fonctionnels simples (contractions globales), l'entraînement en force/endurance ou le retour à l'activité physique normale, ne rétablit pas la coordination des muscles profonds du tronc, ni les modifications neuroplastiques corticales décrites précédemment et associées à la douleur. Or, les preuves s'accumulent à l'effet que l'étendue des changements neuroplastiques corticaux est une caractéristique neurophysiologique clé qui est en corrélation avec le niveau de récupération fonctionnelle suite à des blessures non seulement d'origine vertébrale, mais aussi périphérique et corticale (Boudreau *et al.*, 2010). Les bilans de la littérature confirment donc l'importance de l'utilisation d'une approche incluant un entraînement sensorimoteur fin (Boudreau *et al.*, 2010), car un tel entraînement induit un changement plastique plus important au niveau du cortex moteur qu'un entraînement de force ou d'endurance (Adkins *et al.*, 2006)

Étant donné que l'atrophie des LuM ne se corrige pas naturellement après la résolution d'un premier épisode de lombalgie aiguë (Hides *et al.*, 1996), que la variabilité des APA ne se normalise pas spontanément chez certaines personnes après qu'une douleur expérimentale au dos se soit estompée (Moseley et Hodges, 2006) et que des APA retardés des LuM aient été observés chez les personnes ayant une lombalgie récurrente, et cela en dépit de la résolution des symptômes (MacDonald *et al.*, 2009 ; MacDonald *et al.*, 2010), ces déficiences des muscles profonds semblent de bons marqueurs des mécanismes de récurrence (Hides *et al.*, 1996 ; MacDonald *et al.*, 2009 ; Moseley et Hodges, 2006). En effet, un ECR (Hides *et al.*, 2001) a montré que les exercices de contrôle sensorimoteur, combinés avec une gestion médicale et la reprise de l'activité normale, sont plus efficaces pour réduire les récurrences de maux de dos (taux de récurrence : 35 %) que la gestion médicale et le retour à l'activité (taux de récurrence : 75 %), et cela pour des résultats à long terme (2-3 ans de suivi). Bien que les déficiences des muscles LuM et TrA n'aient pas été évaluées pour corréler les changements possibles avec une amélioration clinique, il s'agit de la seule étude qui s'est intéressée au taux de récurrence.

En ce qui concerne la relation avec la réduction de la douleur et de l'incapacité, les ECR qui comparent les exercices de contrôle sensorimoteur avec toute autre forme de programme d'exercice, toujours chez des patients hétérogènes, conduisent tous à des résultats cliniques équivoques, comme le concluent différents bilans de la littérature (Ferreira *et al.*, 2006 ; Macedo *et al.*, 2009) et une méta-analyse (Smith, B. E. *et al.*, 2014). Cependant, il apparaît que l'approche australienne soit particulièrement bénéfique chez les populations qui ont certaines prédispositions à l'instabilité lombaire (Hides *et al.*, 1996 ; Hides *et al.*, 2001 ; O'Sullivan *et al.*, 1997), ce qui a d'ailleurs justifié le développement de RPC de succès et d'échec pour ce programme d'exercice spécifique (Hicks *et al.*, 2005). Conformément à ce raisonnement, une diminution de l'activation des LuM (mesurée par ultrasons), mais pas de l'activation musculaire des TrA, est associée à la présence des facteurs prédictifs de la RPC de succès de Hicks à un PESL (Hebert *et al.*, 2010). Cela supporte la validité conceptuelle de cette RPC ainsi que l'importance clinique potentielle du muscle LuM comme cible des exercices de stabilisation. Une autre

étude a démontré que l'effet des exercices de contrôle sensorimoteur sur la réduction de la douleur était plus important chez les patients qui avaient une faible capacité à recruter les TrA au départ, comme mesuré par ultrasons (Ferreira *et al.* 2010). En outre, une corrélation modérée ($r = -0,35$) a été obtenue entre l'amélioration du recrutement des TrA, telle que mesurée par ultrasonographie, et la réduction de l'incapacité, à la suite d'un programme d'exercice de contrôle sensorimoteur (Ferreira *et al.*, 2010). De plus, il a été démontré qu'une augmentation de l'épaisseur des TrA et une diminution de l'épaisseur des obliques internes, grâce à différents programmes d'exercice, expliquent jusqu'à 18 % de la réduction de la douleur (Vasseljen et Fladmark, 2010). Malheureusement, aucun ECR n'a comparé directement les approches australiennes et de McGill jusqu'à maintenant.

En résumé, un nombre important d'études soutient l'approche australienne (contrôle sensorimoteur) comparativement à l'approche McGill (force/endurance), même si la plupart des études australiennes ont été réalisées avec de petits échantillons de patients et qu'aucune recherche n'a fourni des preuves directes de la supériorité d'une approche par rapport à l'autre pour réduire la douleur, l'incapacité et les récives. Par contre, une récente revue systématique montre que les exercices réalisés selon l'approche australienne seraient plus efficaces que toutes les autres formes d'exercice pour réduire la douleur et l'incapacité perçue (Ford *et al.*, 2020). Rappelons que les RPC préliminaires (de succès et d'échec du traitement) proposées par le groupe de Hicks (Hicks *et al.*, 2005) sont basées sur l'approche de McGill. Utiliser prématurément l'approche McGill est susceptible de renforcer les patrons défectueux de coordination des muscles profonds qui semblent perpétuer la douleur. Cependant, l'approche de McGill pourrait être plus pertinente après avoir complété l'approche australienne, cela de manière à améliorer l'endurance des muscles du tronc, comme proposé plus tôt (Koumantakis *et al.*, 2005). Fait intéressant, une telle approche « hybride » (australienne suivie par l'approche de McGill) est apparemment ce qui est actuellement utilisé par les physiothérapeutes, du moins au Québec. Par conséquent, si nous voulons être pragmatiques, une RPC développée avec l'utilisation de l'approche hybride est plus susceptible d'être utilisée. Le défi des physiothérapeutes est d'adapter le programme d'exercice à la progression du patient avec l'utilisation de critères appropriés en ce qui a trait à la coordination des muscles profonds du tronc afin de décider quand le patient est prêt à commencer l'approche de McGill.

1.4 Quels mécanismes d'action entrent en jeu ?

« Pour qu'un système de classification soit cliniquement utile, il doit être basé sur l'identification des mécanismes sous-jacents à l'origine du trouble, afin d'orienter les interventions ciblées, qui à leur tour doivent prévoir l'issue du trouble » (O'Sullivan, 2005). Dans cette recherche, les mesures liées à un modèle explicatif de l'adhésion au traitement (modèle d'autorégulation de la maladie de Leventhal) et à deux modèles explicatifs de la chronicité de la douleur, soit un modèle psychologique (Vlaeyen et Linton,

2000) et un modèle neuromusculaire (Panjabi, 1992b), ont été utilisées pour atteindre cet objectif.

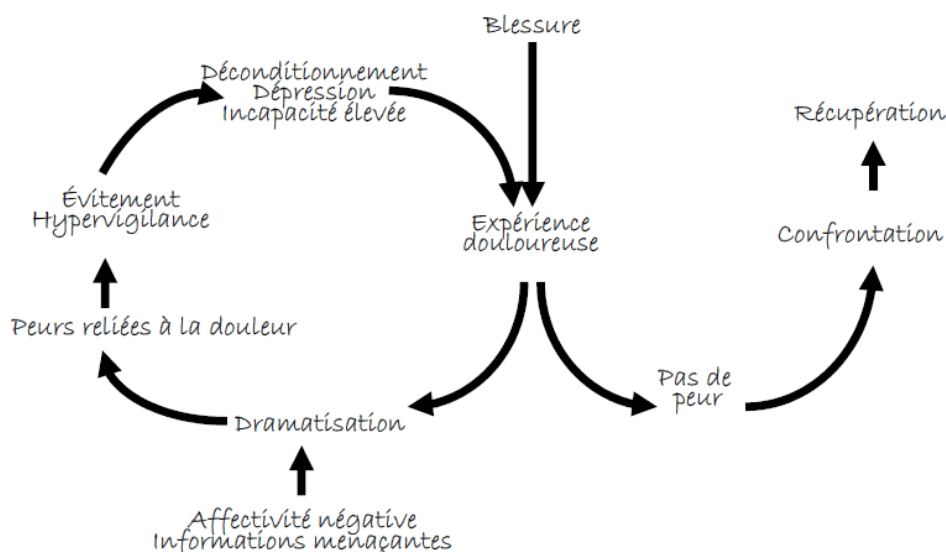
Le modèle d'autorégulation de la maladie ou « *commonsense model of illness* » de Leventhal (Leventhal, 1980 ; Leventhal *et al.*, 2012) est très prometteur pour la recherche et la clinique, notamment pour comprendre l'adhésion au traitement recommandé (Jones *et al.*, 2016 ; L'Heureux *et al.*, 2019). Ce modèle met l'accent sur l'individu, mais tient également compte de sa perception de son environnement. Il s'agit d'un modèle en quatre étapes calquées sur un processus de résolution de problème. Lors de ce processus, les individus recueillent d'abord l'information disponible sur leur état de santé. En fonction de cette information, ils développent une représentation/compréhension personnelle de la condition de santé (Leventhal *et al.*, 2012), également appelée perception de la condition de santé (Foster *et al.*, 2008). Cette représentation est le concept clé du modèle. Elle comprend les dimensions perçues suivantes : causes, conséquences, identité (symptômes et diagnostic perçu), chronologie (aiguë, chronique ou cyclique), représentations émotionnelles (émotions associées à la condition), cohérence (compréhension de la condition de santé) et contrôlabilité (capacités réelles, attentes quant à l'efficacité du traitement et sentiment d'efficacité personnelle) (Leventhal *et al.*, 2012 ; Moss-Morris *et al.*, 2002). Ensuite, en fonction de leur représentation, le patient choisit un plan d'action qui sera en cohérence avec sa compréhension, mais pas nécessairement conforme aux recommandations des intervenants. Cette étape permet donc de comprendre le comportement d'adhésion de l'individu. Finalement, ils évaluent si le comportement amène le changement souhaité par l'individu. Ce modèle considère donc la relation dynamique entre le patient et son environnement. Dans la présente recherche, les éléments de base de ce modèle ont été évalués en utilisant les questionnaires appropriés, notamment en se basant sur différentes revues systématiques portant sur l'adhésion en physiothérapie (Beinart *et al.*, 2013 ; Hall *et al.*, 2010 ; Jack *et al.*, 2010 ; McLean *et al.*, 2010 ; Thompson *et al.*, 2015).

Le modèle psychologique peur/évitement (**figure 1**) est l'un des plus robustes pour expliquer la douleur et l'incapacité à long terme (Smeets *et al.*, 2006 ; Vlaeyen et Linton, 2000). En bref, ce modèle prédit que les patients ayant peu ou pas de craintes face à la douleur confronteraient cette douleur dans leurs activités, ce qui les positionne sur la voie de la récupération. L'autre voie est tout autre, car elle met en action un cercle vicieux. En effet, les patients ayant certaines caractéristiques psychologiques les prédisposant à la dramatisation de la douleur éviteraient l'activité physique (mouvements) qu'ils associeraient à l'exacerbation de la douleur, ce qui conduirait progressivement au déconditionnement physique et à l'émergence de symptômes dépressifs, à la douleur chronique et en définitive à l'incapacité. Heureusement, quelques ECR suggèrent que ces mécanismes sont réversibles avec l'application de différentes modalités d'exercices actifs, menant à une diminution de dramatisation de la douleur (Mannion *et al.*, 1999 ; Marshall *et al.*, 2013 ; Smeets *et al.*, 2006 ; Woby *et al.*, 2004), de la peur du mouvement (Koumantakis *et al.*, 2005), de la peur/évitement (Mannion *et al.*, 1999 ; Storheim *et al.*

2003a ; Woby *et al.* 2004) et des symptômes dépressifs (McCracken *et al.*, 2002). Toutefois, il reste difficile de savoir si les exercices de stabilisation lombaire, qui contiennent moins d'activités physiques vigoureuses que les autres programmes d'exercice, ont des effets sur les variables psychologiques liées à la douleur et l'incapacité. Dans la présente recherche, les éléments de base de ce modèle ont été évalués en utilisant les questionnaires appropriés.

Figure 1. Le modèle peur/évitement.

Adapté de « Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art », par J. W. S. Vlaeyen et S. J. Linton, 2000, *Pain*, 85(3), p. 329. ©Elsevier, 2000.



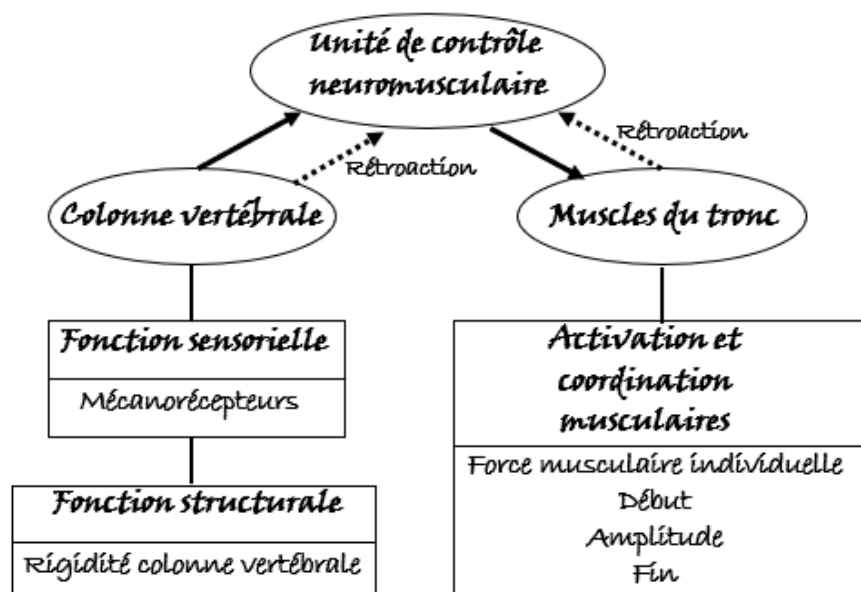
Le modèle de stabilité lombaire (Panjabi, 1992a), qui a été raffiné plus récemment [(Panjabi, 2006), **figure 2**], en est un de type neuromusculaire (biomécanique, physiologique) qui est aussi soutenu par un nombre croissant d'évidences. Ce modèle prédit qu'une lésion d'un tissu passif donné de la colonne vertébrale peut conduire à une instabilité articulaire, laquelle doit être compensée par l'unité de contrôle neuromusculaire (système nerveux) grâce à l'action des muscles entourant la colonne lombaire. Toutefois, des erreurs de contrôle neuromusculaire dues à une proprioception lombaire, à des APA ou à des réponses réflexes déficientes peuvent non seulement être responsables d'une première lésion lombaire, mais peuvent aussi perpétuer le problème, conduisant à une douleur persistante. Or, ces déficits ne sont pas universellement observés chez tous les patients avec lombalgie chronique, certaines études montrant des résultats anormaux et d'autres non. Cela soutient l'idée que ces patients représentent une population hétérogène démontrant diverses déficiences neuromusculaires.

Compte tenu des résultats présentés dans la section 2.2 (Quels mécanismes d'action entrent en jeu ?), un programme d'exercices de stabilisation lombaire devrait impliquer des mécanismes d'action pour normaliser l'atrophie des LuM ainsi que l'inhibition des LuM et TrA. Ce programme devrait également normaliser les APA des LuM et TrA, si ces mécanismes sont déficients. Fait intéressant, la coordination des muscles dorsaux, évaluée en utilisant le phénomène de flexion-relaxation, a été améliorée à la suite d'un programme de stabilisation lombaire (Marshall et Murphy, 2006), contrairement à d'autres formes d'exercices (Mannion *et al.* 2001a). Cependant, d'autres mécanismes d'action n'ont toujours pas été évalués et pourraient mettre en cause la proprioception lombaire, les réponses réflexes et l'équilibre postural du tronc.

Six tests complémentaires (**annexe A.3**) ont été implantés et évalués sur le plan de la validité et/ou de la fidélité test-retest (exception de la proprioception) de manière à couvrir les différentes fonctions neuromusculaires nécessaires pour préserver la stabilité lombaire (Panjabi, 1992b). La fidélité à court terme de la plupart de ces tests a été documentée dans ces études, considérant de 2 jours à 2 semaines d'écart entre les évaluations. Leur fidélité à moyen terme (correspondant à un programme de réadaptation de 8 semaines) a plus récemment été jugée acceptable dans notre étude préliminaire (Larivière *et al.*, 2016), comme résumé dans la section suivante.

Figure 2. Le modèle de stabilisation lombaire.

Adapté de « A hypothesis of chronic back pain: Ligament subfailure injuries lead to muscle control dysfunction », par M. M. Panjabi, 2006, *European Spine Journal*, 15(5), p. 670. ©Springer, 2006.



La stabilisation lombaire est assurée par trois sous-systèmes : la colonne vertébrale (tissus passifs), les muscles du tronc (abdominaux et dorsaux) et l'unité de contrôle neuromusculaire (système nerveux). La colonne vertébrale a deux fonctions : fonction structurale – pour procurer la stabilité mécanique intrinsèque ; fonction sensorielle – pour procurer une rétroaction à l'unité de contrôle sur la posture, le mouvement et le chargement de la colonne vertébrale. L'unité de contrôle coordonne l'activation des muscles du tronc pour assurer la stabilité mécanique de la colonne vertébrale. Il y a aussi rétroaction des muscles du tronc à l'unité de contrôle.

1.5 Résultats de l'étude préliminaire

Les principaux résultats de notre étude préliminaire/faisabilité (Larivière *et al.*, 2016) sont présentés dans cette section. Les trois objectifs étaient : (1) d'amorcer la dérivation des RPC de succès et d'échec du traitement en utilisant des mesures pratiques (cliniques) obtenues lors de l'examen clinique initial ; (2) identifier les mécanismes potentiels associés au succès ou à l'échec du PESL par l'utilisation d'une batterie élargie de mesures psychologiques et neuromusculaires ; (3) évaluer la fidélité test-retest sur 8 semaines (moyen terme) des évaluations neuromusculaires chez des sujets sains n'ayant suivi aucun programme d'exercices.

Le PESL a été mené sur 8 semaines (2 séances/semaine) dans des cliniques de physiothérapie, sans co-intervention. Les patients ont été encouragés à faire les exercices recommandés à domicile. Toutes les mesures sont résumées à l'**annexe A**. En bref, les principales mesures de résultats [douleur (échelle numérique) ; perceptions d'incapacité (échelle Oswestry)] ainsi que plusieurs mesures autorapportées (variables psychologiques ou mesures PSY - associées à la douleur et à l'adhésion au PESL à domicile) ont été recueillies au début (T0), à la quatrième semaine (T4) et à la huitième semaine (T8 - fin) du PESL ainsi qu'à six mois après le traitement. Les autres mesures cliniques qui sont candidates à l'inclusion dans les RPC (**objectif 1**) sont les tests physiques (mesures PHT) faits pendant l'examen clinique, ceci à T0 et T8. Ces tests physiques comprenaient des tests d'instabilité articulaire (n = 4), de flexibilité (n = 6), de déficiences du contrôle sensorimoteur (n = 8), de performance physique (n = 4) et d'endurance musculaire (n = 3). Pour étudier les mécanismes neuromusculaires (mesures NRM ; **objectif 2**) associés au PESL, six tests en laboratoire ont été menés à T0 et T8 chez un sous-échantillon de 32 patients. Ces tests ont été : (1) l'épaisseur et l'activation des muscles profonds du tronc à l'aide de l'imagerie par ultrasons ; (2) la proprioception lombaire ; (3) l'équilibre postural du tronc en position assise sur une chaise instable ; (4) la rigidité lombaire ; (5) les ajustements posturaux anticipatoires (APA) ; (6) la coordination du tronc lors de sa flexion maximale. Afin d'évaluer la fidélité test-retest (**objectif 3**), ces tests ont également été effectués dans le même intervalle de temps (huit semaines) sur un échantillon de 30 sujets sains.

Objectif 1 : Amorcer la **dérivation des RPC** de succès et d'échec au traitement en utilisant des mesures pratiques (cliniques) obtenues lors de l'examen clinique initial. Sur les 64 participants évalués, 48 ont suivi le programme de 8 semaines (attrition de 25 %), dont 31 et 12 patients ont respectivement connu un succès et échec du traitement (voir critères dans la méthodologie (Larivière *et al.*, 2016). Ils ont rapporté une réduction significative de la douleur (d de Cohen = 1,53) et une amélioration de l'incapacité (d = 1,36).

La RPC finale de succès comprenait deux prédicteurs statistiquement significatifs : (1) un test de performance physique (Simmonds, 2006) qui consistait à demander au patient, en position debout, de tenir une légère charge près du corps, à la hauteur des épaules, pour ensuite allonger les bras à l'horizontale vers l'avant jusqu'à la limite de la tolérance, pour en mesurer la distance normalisée à une longueur de bras $> 0,665$ (OR : 13,7 (2,0-94,0)) ; (2) une déficience de contrôle sensorimoteur positive qui était le test d'abduction-rotation passive de la hanche (Van Dillen *et al.*, 1998) (OR : 11,3 (1,6-78,2)). Ensemble, ces deux tests expliquent 39 % de la variance (Nagelkerke R^2). Ce modèle a montré une sensibilité de 94 % (85-100 %) et une spécificité de 50 % (22-78 %), et a démontré un rapport de vraisemblance positif (LR+) de 1,9 (1,0-3,3) et un rapport de vraisemblance négatif (LR-) de 0,13 (0,03-0,55).

La RPC finale d'échec comprenait trois prédicteurs statistiquement significatifs : (1) courbure lombaire $> 29^\circ$ (Ng *et al.*, 2001) (OR : 66,9 (3,5-1268,8)) ; (2) sexe (homme) (OR : 25,5 (1,9-341,6)) et dramatisation de la douleur (OR : 28,8 (2,5-333,9)). Ces trois variables expliquent 62 % de la variance. Ce modèle a montré une sensibilité de 83 % (62-100 %) et une spécificité de 90 % (80-100 %), un LR+ de 8,6 (2,8-26,0) et un LR- de 0,18 (0,05-0,66).

Les prédicteurs identifiés dans ces deux RCP, bien que différents des prédicteurs identifiés dans les RPC préliminaires de Hicks *et al.* (2005), semblent conformes à la théorie qui sous-tend ce programme d'exercices. Veuillez noter le large IC à 95 % pour toutes les statistiques prédictives, ce qui était attendu avec une si petite taille d'échantillon. L'ajout de participants était donc nécessaire pour (1) confirmer l'entrée de ces variables ; (2) considérer de nouvelles variables ; (3) resserrer les intervalles de confiance.

Objectif 2 : Identifier les mécanismes d'action potentiels associés au succès ou à l'échec du PESL grâce à l'utilisation d'une batterie élargie de mesures psychologiques et neuromusculaires, permettant ainsi de décrire les effets spécifiques du traitement.

Les ANOVA et les analyses corrélationnelles ont révélé des effets statistiquement significatifs pour certaines variables de chaque catégorie de mesures (PHT, PSY, NRM) et plusieurs tendances dans les résultats ($0,05 < P < 0,10$), suggérant que le recrutement de participants supplémentaires permettrait de mieux confirmer ou rejeter plusieurs

hypothèses. En effet, la taille des échantillons n'était pas suffisante pour permettre de tester avec une puissance statistique suffisante la présence d'interactions TEMPS (T0 vs T8) × SOUS-GROUPE (succès vs échec), en particulier pour les mesures de NRM qui ont été collectées uniquement dans un sous-échantillon (n = 32) de tous les participants (n = 48). Parmi les mesures NRM, l'imagerie par ultrasons (épaisseur et activation des muscles profonds du tronc) et le test d'équilibre postural en position assise sur une chaise instable ont été les plus sensibles aux changements observés en termes d'incapacité et de douleur au cours du PESL.

Une étude plus approfondie des mesures PSY théoriquement liées à l'adhésion au PESL fait à domicile a été réalisée, car l'adhésion est, à son tour, théoriquement liée à la réussite du traitement. Nos résultats suggéraient qu'il n'était pas possible de prédire l'adhésion au PESL à domicile avec des mesures prises à T0, ce qui pourrait expliquer pourquoi une seule mesure PSY a été retenue dans la RPC d'échec.

Objectif 3 : Évaluer la fidélité test-retest sur 8 semaines (à moyen terme) des évaluations neuromusculaires chez des sujets sains (15 hommes, 15 femmes) qui n'ont suivi aucun programme d'exercice.

Dans l'ensemble, une fidélité modérée à excellente des mesures de NRM a été obtenue, ce qui permet d'avoir une certaine confiance dans ces mesures pour l'étude des différentes fonctions neuromusculaires chez les patients suivant un programme de réadaptation. En effet, cette fidélité était acceptable pour certaines mesures, avec un indice de dépendabilité ID, analogue au coefficient de corrélation intraclasse, supérieur à 0,75, et modéré pour les autres mesures. Quelques mesures ont montré une faible fidélité. Ainsi, pour presque chaque test, il est possible de retenir un sous-ensemble de mesures fidèles.

Deux tests NRM (proprioception lombaire et équilibre postural) ont révélé un certain apprentissage de la tâche entre les deux séances de mesure. Cet apprentissage était négligeable pour le test d'équilibre postural, mais suffisamment important pour rejeter le test de proprioception, qui était un test de perception du mouvement. Pour la suite de cette étude, nous proposons d'utiliser un autre test de proprioception lombaire (sens du repositionnement).

En conclusion, compte tenu des résultats positifs de l'étude préliminaire (Larivière *et al.*, 2016), il avait été recommandé de recruter des participants supplémentaires afin de finaliser la phase de dérivation de ces RPC (**objectif 1**) et mieux faire ressortir les mécanismes PSY et NRM associés au PESL (**objectif 2**).

2. OBJECTIFS DE RECHERCHE

Ce projet poursuit deux objectifs :

(1) **Finaliser la dérivation de la RPC** de succès en utilisant des mesures pratiques (cliniques) obtenues lors d'un examen clinique initial. Les travaux entourant la RPC d'échec ne pouvaient plus être soutenus pour deux raisons. Premièrement, une RPC d'échec a une utilité clinique limitée, comparativement à une RPC de succès. Ceci représente un frein important à l'utilisation d'une RPC d'échec, considérant le temps d'évaluation clinique nécessaire. D'ailleurs, parmi les RPC développées en physiothérapie pour prédire les effets d'un traitement chez des patients avec un trouble musculosquelettique (n = 21 RPC), seulement quelques-unes se sont intéressées à la prédiction de l'échec, comme dans le cas de Hicks *et al.* (2005). Deuxièmement, considérant les trois étapes de développement d'une RPC (dérivation, validation, étude d'impact), comme détaillé ci-dessous, poursuivre le développement d'une RPC d'échec impliquerait de doubler cet effort de recherche. Cela explique pourquoi seulement la RPC de succès de Hicks *et al.* (2005) est passée à l'étape de validation (Rabin *et al.*, 2014). Les efforts ont plutôt été investis vers la prédiction du succès non seulement à la fin du traitement, mais également au suivi de six mois après le traitement, comme le suggèrent les auteurs d'une revue systématique sur le sujet (Beneciuk *et al.*, 2009), ce qui en fait la première étude sur le sujet.

Rappelons que le développement d'une RPC se fait en trois étapes (Childs et Cleland, 2006) :

1. Dérivation : Pour établir les facteurs ayant une valeur prédictive ;
2. Validation : Pour évaluer l'efficacité de la RPC avec une population différente de patients ou avec un établissement de soins différent, par exemple avec d'autres physiothérapeutes rattachés à différentes cliniques. Cela nécessite un ECR afin de s'assurer que les prédicteurs déterminés lors de l'étape de dérivation ne sont pas apparus par hasard ;
3. Étude d'impact : Pour fournir la preuve que la RPC change le comportement des cliniciens et améliore les résultats cliniques des patients et/ou réduit les coûts. Différents plans d'étude peuvent être utilisés, dont une analyse coût/bénéfice.

(2) **Établir les mécanismes d'action** impliqués dans ce programme d'exercice à l'aide d'une batterie étendue de mesures, permettant ainsi aux effets spécifiques du traitement d'être décrits. En d'autres termes, nous voulons expliquer les raisons pour lesquelles certains patients ont une réponse positive aux exercices de stabilisation lombaire et d'autres pas. Dans cet ordre d'idées, une attention particulière a été portée à l'adhésion au programme d'exercice à domicile.

Hypothèses de recherche :

(H1) En lien avec l'objectif 1 : Un sous-groupe de patients présentera une nette amélioration clinique, comme documenté par le questionnaire d'Oswestry (OSW) portant sur l'incapacité perçue, et un autre sous-groupe présentera un échec clinique. Ces sous-groupes seront caractérisés par des mesures physiques et psychologiques spécifiques, telles que mesurées à l'examen initial.

(H2) En lien avec l'objectif 2 : Le sous-groupe de patients présentant une nette amélioration clinique démontrera des progrès dans le temps tels que documentés avec les mesures physiques (PHT), psychologiques (PSY) et neuromusculaires (NRM), alors qu'aucune différence ne sera observée dans le sous-groupe démontrant un échec clinique, menant à des interactions SOUS-GROUPE × TEMPS statistiquement significatives.

(H3) En lien avec l'objectif 2, mais plus spécifiquement, avec les mesures PSY : Différentes variables mesurées par questionnaires permettront d'expliquer l'adhésion au programme d'exercice fait à domicile.

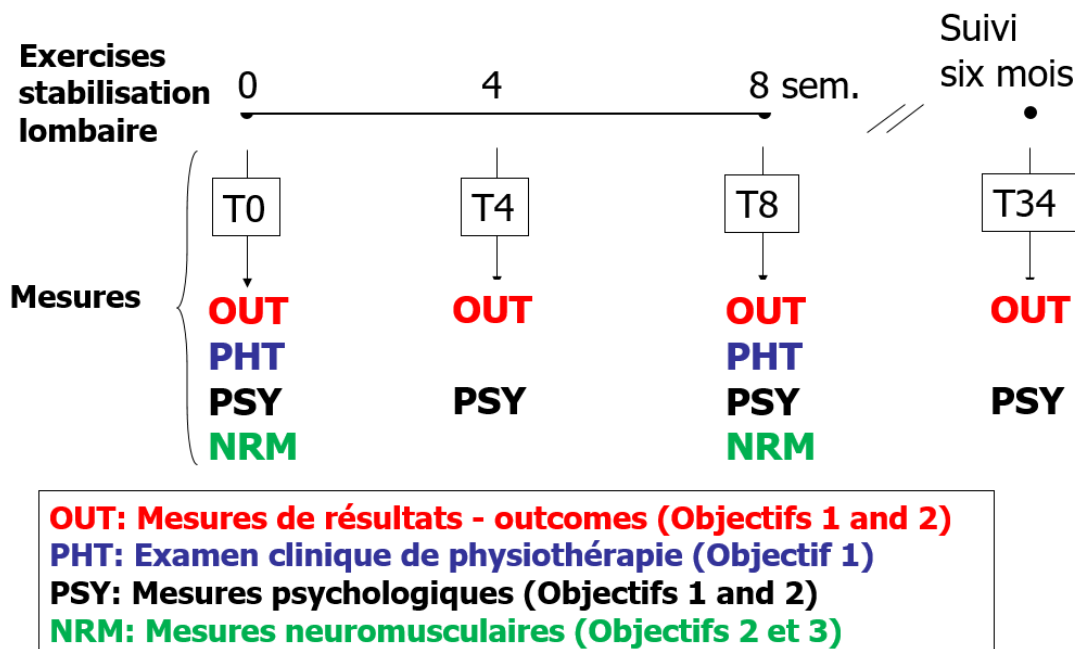
3. MÉTHODOLOGIE

3.1 Devis expérimental

Comme illustré à la **figure 3**, nous avons réalisé une étude prospective comptant trois évaluations au cours de l'intervention, ainsi qu'une quatrième évaluation au suivi de six mois chez un seul groupe de patients. Afin de réduire le risque que les RPC retiennent des mesures non spécifiques au traitement ou des mesures qui ressortent de façon erronée (erreur de type 2), certains méthodologistes suggèrent l'utilisation d'un ECR avec groupe contrôle pour la dérivation de la RPC (Fitzgerald, 2010 ; Stanton *et al.*, 2010). Cependant, toutes les RPC à ce jour ont utilisé un devis avec un seul groupe à l'étape de dérivation (Stanton *et al.*, 2010). Nous avons également choisi cette stratégie et prévoyons de passer le test de l'ECR uniquement à l'étape de validation des RPC. Cependant, afin d'augmenter les chances de passer l'étape de validation avec succès, seules les mesures disposant d'un lien conceptuel avec la stabilité lombaire ou avec l'adhésion au programme d'exercice ont été retenues comme candidates.

Ainsi, pour la dérivation de la RPC (**Objectif 1**), les patients ont été évalués par des mesures cliniques (examen clinique et questionnaires), au début du programme (T0), à la fin (T8), et six mois plus tard (T34). Pour étudier les mécanismes d'action impliqués dans ce programme d'exercice spécifique (**Objectif 2**), les patients recrutés pour l'objectif 1 ont également été évalués à mi-parcours du programme d'exercice (T4) au regard de mesures psychologiques, et à T0 et T8, au regard de l'évaluation neuromusculaire en laboratoire. Des mesures neuromusculaires (NRM) ont déjà été recueillies (étude préliminaire) chez 30 sujets sains à T0 et T8 (sans traitement) pour les comparer aux patients à T0 et évaluer leur fiabilité test-retest à 8 semaines (Larivière *et al.*, 2016).

Figure 3. Quatre catégories de mesures (OUT, PHT, PSY, NRM) ont été recueillies chez les patients à différents moments (T0, T4, T8) et après (T34) le programme d'exercice. Les mesures PHT et NRM sont plus coûteuses et n'ont donc été collectées qu'à T0 et T8.



3.2 Participants

La taille de l'échantillon nécessaire pour répondre à l'objectif 1 (dérivation de la RPC) a été estimée selon deux méthodes à partir des données préliminaires (Larivière *et al.*, 2016). **À noter qu'elle avait été déterminée à l'époque en vue du développement des deux RPC (succès et échec) de l'effet post-traitement (T8 seulement), ce qui explique les prochaines estimations prenant aussi en compte la RPC d'échec.** Dans la première méthode, une analyse de puissance a priori a été réalisée à l'aide du programme GPower 3.1 (Faul *et al.*, 2009), en utilisant la procédure proposée par Demidenko (2007) pour une estimation de la taille de l'échantillon spécifique à la régression logistique. Cette procédure est considérée comme plus précise que les autres (Hsieh, 1989 ; Lyles *et al.*, 2007). Sur la base des données de notre étude préliminaire (N=43 ; succès : n = 31 ; échec : n = 12), deux modèles (succès/échec) ont été construits. La taille d'effet pour la **RPC de succès** a été estimée en considérant un rapport de cote de 13,7 (95 % IC 2,0-94,0). En ajustant pour les autres covariables du modèle en utilisant un facteur d'inflation de la variance estimé dans l'étude préliminaire, et en considérant un alpha de 0,05 et une puissance de 0,80, la taille de l'échantillon projeté est de 39 patients ; où 49 patients en tenant compte d'un taux d'attrition de 25 % (16/64 de l'étude préliminaire).

Ainsi, 18 (= 49-31) nouveaux patients ont été recrutés pour atteindre la puissance requise pour le modèle de réussite de l'étude, comme détaillé dans le **tableau 1**.

Pour la RPC d'échec, la taille de l'effet (ou rapport de cote) était de 19,4 (95 % IC 2,4-157,0), ce qui donne 35 patients ou 44 en ajustant pour les autres covariables du modèle et le même alpha, la même puissance et la même attrition. Étant donné que 12 patients ont déjà été recrutés, il faudrait en recruter 32 autres (=44-12) pour le modèle d'échec. Cela signifie qu'il faudrait 50 nouveaux patients (18 + 32) pour atteindre un échantillon total de 93 patients en combinant les études préliminaire (n = 49) et actuelle (n = 44), comme défini dans le **tableau 1**.

Tableau 1. Première méthode d'estimation de la taille d'échantillon

	RPC de succès	RPC d'échec	Total
Étude préliminaire	31	12	43
Présente étude	18	32	50
Taille d'échantillon ciblée	49	44	93

La deuxième méthode conduit aux estimations détaillées dans le **tableau 2**. Cette méthode considère simplement que les effectifs des deux groupes à recruter seront inégaux. Le rapport entre les nombres déjà recrutés $k = 31/12 = 2,6$ suggère que pour chaque 2,6 avec succès il y aura 1 échec. Il est donc possible d'ajuster la taille initialement calculée par ce facteur selon l'équation suivante : $N = n (1 + k)^2 / 4k$.

Tableau 2. Deuxième méthode d'estimation de la taille d'échantillon

	RPC de succès	RPC d'échec	Total
Étude préliminaire	31	12	43
Présente étude	18	46	64
Taille d'échantillon ciblée	49	58	107

Considérant que n est égal à 50 patients, c'est-à-dire le nombre déjà calculé par l'étude de puissance, et que K est le facteur d'ajustement = 2,6, cette équation conduit à une estimation de 64 patients au lieu des 50 initialement calculés avant de prendre en compte le différentiel de recrutement entre les deux groupes.

Deux autres équations permettent de répartir ces 64 patients dans les 2 RPC :

Pour la RPC de succès : $N_s = N / (1 + k) = 63 / (1 + 2,6) = 63 / 3,6 = 17,5$; soit 18 patients comme estimé ci-dessus.

Pour la RPC d'échec : $N_e = KN / (1 + k) = 2,6 \times 63 / (1 + 2,6) = 163,8 / 3,6 = 45,5$; soit 46 patients.

Pour atteindre l'objectif 2 (mécanismes), en utilisant les taux de prévalence obtenus à partir de l'étude préliminaire (succès : $31/48 = 0,65$; échec : $12/48 = 0,25$), nous estimons

que 64 nouveaux patients conduiraient à un total de 73 (31 anciens + 42 nouveaux) et 28 (12 anciens + 16 nouveaux) patients dans les sous-groupes succès et échec, respectivement, ce qui est suffisant pour étudier les mécanismes mesurés avec les variables PHT et PSY, en utilisant des ANOVA (détails dans la section des analyses statistiques).

Dans la présente étude, un échantillon de **64 patients (32 hommes et 32 femmes)** était donc visé, en recrutant à la clinique externe de l'Institut de réadaptation Gingras-Lindsay-de-Montréal (IRGLM), dans certaines cliniques de physiothérapie de la région métropolitaine et par le biais d'annonces dans les journaux. Les critères d'inclusion étaient les suivants : (1) être âgé entre 18 et 65 ans ; (2) maîtrise de la langue française ou anglaise ; (3) présence de douleur lombaire ou lombo-sacrée non aiguë, définie comme une douleur quotidienne ou quasi quotidienne depuis plus de 4 semaines, avec ou sans douleur radiculaire ; (4) obtenir un score supérieur à 12 % au questionnaire Oswestry (OSW), le changement minimal cliniquement important étant de 10 points (voir la section des analyses statistiques). Les principaux critères d'exclusion étaient les suivants : (1) pathologie lombaire spécifique (fracture, infection ou tumeur) ou scoliose ; (2) signe neurologique positif dans deux des trois catégories de test (1. réflexes au genou et la cheville; 2. myotomes; 3. dermatomes) ; (3) chirurgie de l'appareil musculosquelettique du tronc ; (4) maladie systémique ; (5) début d'un programme d'exercice durant les derniers six mois ; (6) grossesse ; (7) douleur dorsale ou cervicale plus intense que la douleur lombaire ; (8) litige sur le plan juridique lié à la blessure au dos. Les patients non obèses (indice de masse corporelle > 30 kg/m²) ont été sollicités pour les mesures NRM en laboratoire.

Une approche per protocole a été adoptée au lieu d'une approche en intention de traiter (*intention to treat*), car elle est plus pertinente à l'étape de dérivation d'une RPC. Un critère d'exclusion additionnel était donc une assiduité insuffisante aux rendez-vous cliniques (absences sans prévenir à deux séances consécutives). Seulement l'adhésion au programme d'exercice à domicile était laissée au choix des patients, car cela correspond à la pratique.

Les collectes de données ont été réalisées à l'IRGLM, soit un établissement du Centre de recherche interdisciplinaire en réadaptation (CRIR) du Montréal métropolitain. Avant de participer, les participants ont donné leur consentement éclairé par écrit. Les droits des participants ont été protégés. Toutes les procédures ainsi que le formulaire de consentement ont été approuvés par le comité d'éthique du CRIR (numéros de certificats d'éthique : CRIR-738-0512 et CRIR-1315-0318 pour les études préliminaire et finale, respectivement).

3.3 Intervention

Le programme d'exercice de stabilisation lombaire de 8 semaines a été pris en charge par 11 physiothérapeutes de cliniques privées de physiothérapie de la région de Montréal, sans co-interventions d'autorisés (ex. : ergothérapie, TENS, massage, ultrasons, etc.), sauf pour les médicaments (anti-inflammatoires non stéroïdiens, analgésiques, opioïdes, relaxants musculaires) qui ont été enregistrés. Afin d'uniformiser l'information donnée aux patients sur leur état, une brochure intitulée *Le Guide du dos* (Roland *et al.* 2009) a été remise aux patients à leur première visite clinique. L'objectif principal de cette brochure est de changer les croyances et les comportements (reprise des activités) relatifs aux maux de dos. Il s'agit de la traduction en français du *Back Book* (Burton *et al.*, 1999). Les physiothérapeutes ont reçu une formation d'une journée portant sur la description de l'étude (1 heure) et surtout, sur la façon d'enseigner les exercices aux participants (8 heures).

Le programme d'exercice, qui a été détaillé antérieurement (Larivière *et al.*, 2016; consulter l'**annexe A**), respectait essentiellement l'approche australienne (Richardson *et al.*, 2004), qui se concentre sur le contrôle sensorimoteur des muscles profonds du tronc, suivie par l'approche de McGill (2002), comprenant des exercices de surcharge. Il se déclinait en trois phases, soient (1) la gestion de la douleur et le contrôle sensorimoteur isolé des muscles profonds ; (2) le niveau débutant de contrôle des déficiences et fonctions (initiation des exercices avec accent sur la qualité du contrôle du mouvement) ; (3) le niveau modéré/avancé de contrôle des déficiences et fonctions (développement de l'endurance avec accent sur quantité).

L'utilisation de méthodes de rétroaction plus accessibles (excluant l'imagerie par ultrasons) telles que la palpation et l'unité de rétroaction de pression (Richardson *et al.*, 2004) a été privilégiée dans la présente étude afin de mieux généraliser nos résultats à la pratique actuelle en physiothérapie. Les physiothérapeutes ajustaient la progression selon les besoins des patients. Tous les participants ont fait le programme en clinique à raison de deux séances de 30 min par semaine sur une période de 8 semaines (sauf première visite de 60 min pour fins d'évaluation par le physiothérapeute attitré; aucune information n'a été tirée de cette évaluation par l'équipe de recherche). Pendant et après le programme de 8 semaines, les patients étaient encouragés à faire les exercices à domicile, à l'aide d'une feuille indiquant le type d'exercice et ses paramètres d'exécution, selon le calendrier suivant : 7, 3 et 3 jours par semaine pour les phases 1, 2 et 3 du programme d'exercices, respectivement. Après le programme clinique, les patients étaient libres d'avoir recours à tout type de soin, ce qui correspond à une approche plus pragmatique, mais comporte également des limites au niveau de la dérivation de la RPC de succès au suivi de six mois.

3.4 Évaluations

Pour éviter toute interférence dans la relation entre le patient et son physiothérapeute, toutes les évaluations ont été effectuées par le personnel de recherche, en marge de l'intervention. De plus, il a été précisé aux patients et physiothérapeutes que les résultats de ces évaluations ne leur seraient pas transmis. En plus des caractéristiques démographiques de base (âge, sexe, état civil, scolarité) et certains éléments de l'histoire médicale de la personne, les quatre catégories de mesures identifiées à la **figure 3** et présentées plus en détail dans les prochaines sections (OUT, PHT, PSY, NRM) ont été considérées (détails additionnels à l'**annexe A**). Pour éviter les biais de mémoire, la durée de la douleur a été évaluée avec une question (**Depuis combien de temps souffrez-vous de douleurs au dos ?**) et un choix de réponses : (1) moins de 1 mois ; (2) 1-3 mois ; (3) 3-6 mois ; (4) 6 mois-1 an ; (5) 1-5 ans ; (6) plus de 5 ans.

3.4.1 Mesures de résultats (OUT)

Quatre mesures de résultats ont été prises en compte : perception de l'incapacité, intensité de la douleur, le changement global (ChangGlob) et la satisfaction du patient (**annexe A.1**). Le questionnaire Oswestry (**OSW**) sur la perception de l'incapacité liée à une lombalgie (Fairbank *et al.*, 1980) a été choisi comme la mesure critère pour développer la RPC de succès, comme c'est la pratique pour le développement de RPC dans le domaine de la physiothérapie.

3.4.2 Mesures cliniques de physiothérapie (PHT)

L'examen physique comprenait des tests qui ont été faits à T0 et T8 et couvraient différentes dimensions. Ces tests devaient démontrer une fidélité interjuges acceptable ($\kappa > 0,6$); coefficients de corrélation intraclasse [$\text{CCI} > 0,75$]), comme détaillé à l'**annexe A.2**. Il y avait d'abord des **tests qui sont théoriquement directement liés au concept de stabilité lombaire** (Alqarni *et al.*, 2011; Ferrari *et al.*, 2015; Hicks *et al.*, 2003), tels que différents tests manuels d'instabilité ou la présence de mouvements aberrants ou de laxité ligamentaire généralisée. Des **tests de déficience du contrôle sensorimoteur** ($n = 7$) ont également été considérés, car la stabilité lombaire serait intrinsèquement liée au contrôle sensorimoteur. Par contre, il a été choisi de n'évaluer que les symptômes provoqués par ces tests, car les compétences nécessaires pour l'évaluation des signes (alignement, mouvements) apparaissent beaucoup plus difficiles à acquérir (Van Dillen *et al.*, 1998), ce qui rendrait l'application d'une éventuelle RPC plus complexe pour la majorité des cliniciens. Les tests proposés par Luomajoki *et al.*, (2007), dont certains ont démontré une bonne fidélité interjuges ($\kappa > 0,6$), n'ont donc pas été retenus. Des **tests d'endurance des muscles du tronc** ($n = 4$) et de **performance physique** ($n = 4$) ont aussi été retenus, car la condition musculaire apparaît aussi liée à la stabilité lombaire. La **posture lombaire** (inclinaison bassin, courbure lombaire) a été évaluée, car des simulations supportent une association positive avec l'instabilité lombaire

(Kiefer *et al.*, 1997) et parce que les muscles LuM contrôlent la courbure lombaire (Claus *et al.*, 2009). Finalement, dans un but stratégique, des tests d'**amplitude de mouvement** (n = 5) ont été introduits, car à la connaissance des auteurs, il y a absence de résultat établissant une relation entre la mobilité et l'instabilité lombaires. En effet, l'absence de leur contribution dans la RPC démontrera que cette dernière ne sera pas en lien avec des mécanismes d'action qui pourraient être associés à n'importe quel programme d'exercices.

Des décisions ont dû être prises afin de gérer les mesures des côtés gauche et droit (ex. : flexion latérale droite et gauche du tronc ou mesures sur les membres inférieurs gauche et droit), car les déficiences peuvent souvent être unilatérales. Conséquemment, au lieu de faire la moyenne des mesures des deux côtés, la valeur minimale a été retenue pour les mesures d'étendue de mouvement et la mesure d'endurance (variable TEM-lat) alors que la valeur maximale a été retenue pour le test de performance physique (variable TPP-rouler), le but étant de retenir les mesures les plus associées à une déficience. En ce qui a trait aux mesures de déficience du contrôle sensorimoteur, menant à trois réponses possibles (-1 = réduction douleur; 0 = aucun effet; +1 = hausse de la douleur), la valeur maximale des deux côtés a été retenue.

3.4.3 Mesures psychologiques (PSY)

Les mesures psychologiques ont d'abord été sélectionnées en fonction du modèle peur/évitement afin d'investiguer les mécanismes d'ordre psychologique (**objectif 2**). Or, dans la poursuite de l'**objectif 1** (RPC), d'autres mesures psychologiques ont été choisies en lien avec l'adhésion au programme d'exercice. En effet, en plus des mesures associées au modèle peur/évitement (Vlaeyen et Linton, 2000), des mesures en lien avec le modèle d'autorégulation de la maladie (Leventhal, 1980) ont été considérées, cela en se basant sur différentes revues systématiques portant sur l'adhésion en physiothérapie (Beinart *et al.*, 2013; Hall *et al.*, 2010; Jack *et al.*, 2010; McLean *et al.*, 2010; Thompson *et al.*, 2015).

Les variables du modèle peur/évitement ont été évaluées avec des questionnaires autoadministrés qui ont des propriétés psychométriques satisfaisantes (**annexe A.1**). Outre les variables de douleur et incapacité déjà mesurées, les variables importantes du modèle peur/évitement sont la dramatisation de la douleur (Sullivan *et al.*, 1995), les peurs et croyances (Waddell *et al.*, 1993), et la détresse psychologique (Prévile *et al.*, 1992). Une mesure de la détresse psychologique a été préférée à celle de la dépression pour obtenir une plus grande variabilité intersujet. Le questionnaire STarT Back (9 items) a également été utilisé comme outil de dépistage (Hill *et al.*, 2008) pour évaluer s'il est possible de prédire les échecs thérapeutiques en fonction du profil psychologique.

Telles que détaillées à l'**annexe A.1** (questionnaires utilisés), différentes variables (facteurs individuels, barrières dans l'environnement, relation thérapeute/patient) ont servi

à prédire l'adhésion au programme d'exercice à domicile (et indirectement le succès du traitement avec la RPC). Les facteurs individuels ont été considérés avec les variables clés du modèle d'autorégulation de la maladie de Leventhal (1980). Il s'agit d'abord de la perception de la maladie (Broadbent *et al.*, 2006) et de la représentation émotive par la détresse psychologique (Préville *et al.*, 1992). Conformément au concept de représentation, plus spécifiquement à notre problématique, la perception des symptômes a été mesurée par la douleur perçue durant l'exercice (Damsgard *et al.*, 2010). Pour mesurer la contrôlabilité, la crédibilité et les attentes du traitement (Deville et Borkovec, 2000), les attentes du patient sur le traitement par l'exercice (Resnick *et al.*, 2000) de même que le niveau d'activité physique effectué durant les semaines précédentes (Baecke *et al.*, 1982) ont été mesurés. La mesure de l'adhésion est déjà prévue comme variable secondaire et correspond au comportement/plan d'action de l'individu. Finalement, l'évaluation du comportement par l'individu a été documentée par la perception des changements globaux. Parallèlement au modèle d'autorégulation de la maladie, les variables relatives aux obstacles à l'adhésion (Jack *et al.*, 2010) ont été mesurées. Ces variables mesurent des concepts complémentaires associés à l'adhésion aux traitements thérapeutiques (Brawley et Culos-Reed, 2000; Coutu *et al.*, 2000). Il s'agit des barrières dans l'environnement, mesurées par une échelle portant sur le sentiment d'efficacité personnelle pour l'exercice (McAuley, 1992) et une autre mesurant le soutien social ou familial à l'activité (Sallis *et al.*, 1987). Finalement, l'alliance thérapeutique entre le patient et son thérapeute, en lien avec le concept d'alliance de travail (Bordin, 1979), a également été mesurée (Corbière *et al.*, 2006; Horvath et Greenberg, 1989). Deux de ces mesures PSY (la crédibilité et les attentes du traitement; l'alliance thérapeutique entre le patient et son thérapeute) ont été recueillies après la première semaine (à T1 au lieu de T0). Celles-ci dépendent de la description du programme d'exercice faite par le physiothérapeute et elles nécessitent un minimum d'interactions entre le patient et le thérapeute pour être valides. Tous les questionnaires étaient disponibles en versions française et anglaise. Comme détaillé à l'**annexe A.1**, la plupart de ces variables figurent parmi les candidates pour la dérivation de la RPC (**objectif 1**) et pour évaluer les mécanismes d'origine psychologique (**objectif 2**).

L'adhésion au programme d'exercice à domicile n'a pas été évaluée pendant la phase clinique de huit semaines pour éviter de modifier le comportement du patient (Jordan *et al.*, 2010) et ainsi de biaiser la dérivation de la RPC. Or, il n'existe aucun standard pour mesurer l'adhésion (Beinart *et al.*, 2013). Aux temps T8 et T34, en lien avec la définition de l'adhésion (Haynes *et al.*, 1979) nous avons posé la question suivante : combien de fois avez-vous fait vos exercices comme prescrit durant la dernière semaine ? Une récente revue systématique a justement permis de constater que la fréquence est la principale mesure utilisée dans l'évaluation de l'adhésion aux exercices (Beinart *et al.*, 2013). Étant donné que l'adhésion est le degré avec lequel le comportement de la personne concorde avec les recommandations du professionnel de la santé, le ratio entre la fréquence des exercices et la fréquence prescrite par le thérapeute a été calculé. Les ratios au-dessus de 1 ont été tronqués à 1 (Mannion *et al.*, 2009).

3.4.4 Mesures neuromusculaires (NRM)

3.4.4.1 Techniques de mesure

Les instruments de mesure utilisés seuls ou en combinaison avec un des montages, comme illustré à l'**annexe A.3**, sont de différents ordres.

Imagerie par ultrasonographie. Un appareil à ultrasons de la compagnie Philips (Philips Medical Systems, Bothell, WA; modèle HD11 1.0.6) a été utilisé, avec une sonde linéaire de 50 mm (modèle L12-5, 12-5 MHz) pour les abdominaux et une sonde curvilinéaire (75°, 6,5 cm; modèle C5-2, 5-2 MHz) pour les LuM.

Électromyographie de surface. Pour les tests 5 (ajustements posturaux anticipatoires) et 6 (coordination du tronc) décrits ci-dessous, les signaux EMG des muscles dorsaux (n = 6), abdominaux (n = 4) et de l'épaule droite ont été recueillis (taux d'échantillonnage : 1000 Hz; bande passante : 20-450 Hz) avec un système Bagnoli-16 (DS-B04; Delsys Inc., MA) et 11 paires d'électrodes de surface (modèle DE-2.1; gain : 10, Delsys Inc., Wellesley, MA). Les électrodes ont été positionnées bilatéralement vis-à-vis le muscle multifide au niveau L5, l'ilio-costal lombaire au niveau L3, et le long dorsal à L1, de même que sur les obliques externes (OE), les obliques internes (OI) et le deltoïde antérieur.

Cinématique. La cinématique angulaire de la chaise instable (test 3), du tronc et du bras droit a été mesurée avec un système comprenant cinq senseurs 3D (X-Sens Motion Technologies, Enschede, Hollande) positionnés sur la chaise, le sacrum, les vertèbres T12 et C4, ainsi que sur le bras droit. Ces mesures étaient nécessaires pour les tests 3 (équilibre postural), 5 (ajustements posturaux anticipatoires) et 6 (coordination du tronc) décrits ci-dessous. Les senseurs positionnés sur le tronc et le bras ont été installés seulement après le test 3.

3.4.4.2 Tests neuromusculaires

Six tests ont permis d'évaluer les différentes fonctions neuromusculaires potentiellement associées à la stabilité lombaire (détails et figures à l'**annexe A.3**), comme suggéré par le modèle de la stabilité lombaire (Panjabi, 1992b). La plupart de ces tests ont été implantés et évalués au regard de leur validité ou fidélité test-retest à court terme (2 semaines d'écart), et à moyen terme (8 semaines d'écart).

Test n° 1. Épaisseur et activation des muscles profonds du tronc avec l'imagerie par ultrasons. L'épaisseur des muscles LuM (muscles profonds du dos à L3, L4 et L5) et des abdominaux (OE, OI, TrA) a été mesurée avec les images correspondant au repos et à l'activation musculaire, de manière à pouvoir détecter une possible atrophie musculaire (mesure au repos) ou une activation non optimale de ces muscles, en calculant le pourcentage de changement entre l'état de contraction et l'état de repos (Larivière *et al.*, 2018; Larivière, Henry, *et al.*, 2019). À la lumière des résultats obtenus dans l'étude

préliminaire (Larivière, Henry et Preuss, 2020; Larivière, Preuss, *et al.*, 2020), d'autres mesures ont été considérées. L'épaisseur du fascia thoracolombaire et des fascias qui séparent les muscles abdominaux a également été quantifiée, car ces fascias joueraient un rôle dans le contrôle de la stabilité lombaire. Finalement, une zone (aire) d'intérêt a été tracée automatiquement sur les LuM (à L3, L4 et L5) pour calculer son niveau d'échogénicité (en fonction des niveaux de gris des pixels dans la zone d'intérêt), laquelle étant liée aux infiltrations graisseuses ou formation de fibrose intramusculaire, soit un indicateur de sous-utilisation des LuM.

Test n° 2. Proprioception lombaire. Le test permet de quantifier le sens du repositionnement en mode actif, avec les yeux fermés. Il s'agit de l'erreur entre : (1) une position cible de la chaise (10° à gauche ou à droite), telle que générée par un moteur et ensuite mémorisée par le sujet ; (2) la position que le sujet atteint, par l'action de ses muscles du tronc, pour s'approcher le plus près possible de cette position cible mémorisée (Boucher *et al.*, 2017; Lee, A. S. *et al.*, 2010). Note de rappel : le test réalisé dans l'étude préliminaire évaluait plutôt le sens du mouvement, mesuré avec le même appareil, mais menant à des résultats équivoques (Boucher *et al.*, 2016). Il ne sera donc pas possible d'utiliser ces données dans la présente étude.

Test n° 3. Équilibre postural du tronc en position assise sur une chaise instable (yeux fermés). Un capteur collé sur l'assise de la chaise, derrière le sujet, mesure les mouvements avant/arrière et médio-latéral de la chaise, ce qui permet de générer une représentation 2D du mouvement de la chaise, appelée stabilogramme. Ce stabilogramme est quantifié de différentes façons, notamment avec des variables permettant de connaître l'amplitude des mouvements, mais aussi le contenu fréquentiel de ces mouvements. Vingt-quatre variables ont été retenues, parmi les 38 calculées antérieurement, sur la base des résultats de fidélité et de sensibilité au PESL (Larivière, Mecheri, *et al.*, 2013; Shahvarpour *et al.*, 2018).

Test n° 4. Rigidité lombaire (rigidités intrinsèque et réflexive). Les signaux de force et de mouvement du tronc (du harnais entourant le thorax) qui ont été recueillis lors des perturbations avant/arrière du tronc/harnais ont été utilisés dans un modèle mathématique permettant de départager les contributions intrinsèques et réflexives responsables de la rigidité lombaire, menant à l'estimation de 12 variables (Larivière *et al.*, 2015). Il est à noter que le test a été légèrement modifié dans le but d'obtenir des résultats plus probants que ce qui a été obtenu dans l'étude préliminaire (Larivière *et al.*, 2016). En résumé, au lieu de faire des essais avec une préactivation des muscles du dos équivalant à la production d'un moment de force en extension de 10 Nm (femmes) et 15 Nm (hommes), les essais ont été répartis avec différentes intensités de préactivation (femmes : 0, 4, 8 et 12 Nm; hommes : 0, 6, 12 et 18 Nm) afin d'établir la relation linéaire entre ces intensités et les 12 variables initiales. Des analyses indépendantes ont donc été réalisées pour les 42 patients qui ont suivi ce nouveau protocole. Afin de faire des analyses incluant les participants de l'étude préliminaire, il a été possible d'utiliser cette

relation linéaire pour estimer (interpoler) les 12 variables pour des intensités de 10 Nm (femmes) et 15 Nm (hommes).

Test n° 5. Ajustements posturaux anticipatoires (APA). Il s'agit d'abord de quantifier la latence entre l'activation du muscle deltoïde antérieur (épaule) droit et l'activation de chacun des muscles du tronc, ce qui crée 10 variables de latence (électrodes EMG sur 10 muscles du tronc) (Boucher *et al.*, 2018; Larivière, Butler, *et al.*, 2013). Plusieurs de ces latences sont négatives, indiquant que certains muscles du tronc s'activent avant l'initiation du mouvement du bras (mesuré avec l'EMG du deltoïde antérieur), cela afin de contrer la perturbation imminente du tronc. Il s'agit ensuite de vérifier si cette préactivation des muscles du tronc est efficace pour limiter les mouvements de la colonne lombaire tels que mesurés avec les senseurs de mouvement. Le mouvement de la colonne lombaire a donc aussi été quantifié en fonction de la position (Boucher *et al.*, 2018; Larivière, Butler, *et al.*, 2013), de la vitesse et de l'accélération angulaires maximales enregistrées entre le moment précédant (- 100 ms) et suivant (200 ms) l'activation du deltoïde antérieur.

Test n° 6. La coordination du tronc. Cette coordination a d'abord été quantifiée avec l'EMG de surface des muscles dorsaux, c.-à-d. en quantifiant le phénomène de flexion-relaxation (Shahvarpour *et al.*, 2017):

$$\text{Ratio flexion/relaxation (RFR)} = \frac{\text{EMG maximal lors du mouvement de flexion}}{\text{EMG minimal en fin de flexion}}$$

L'étendue de mouvement des trois segments du tronc (bassin, colonnes lombaire et thoracique) et du tronc en entier a aussi été quantifiée avec des senseurs de mouvement positionnés sur le sacrum (bassin) et sur la colonne vertébrale en deux endroits (T12 et C7), permettant d'évaluer les contributions de chaque segment. Afin de quantifier la coordination entre les segments, la contribution relative de chacun de ces segments au mouvement maximal du tronc (senseur à C7) a été calculée (Shahvarpour *et al.*, 2017).

3.5 Analyses statistiques

3.5.1 Objectif 1 – Dérivation de la RPC de succès (hypothèse 1)

Les procédures statistiques généralement utilisées pour développer une RPC ont été utilisées (Cleland *et al.*, 2007; Flynn *et al.*, 2002; Hicks *et al.*, 2005). Tout d'abord, trois sous-groupes de patients ont été définis en fonction de leur niveau de succès à l'aide de l'OSW à T8 : (1) succès ; (2) amélioration cliniquement significative ; (3) échec, de façon similaire à ce que propose Hicks *et al.* (2005). Pour ce faire, le score de changement de l'OSW ($\Delta\text{OSW} = \text{OSWT0} - \text{OSWT8}$) et le pourcentage correspondant [$\Delta\text{OSW} \% = ((\text{OSWT0} - \text{OSWT8}) / \text{OSWT0}) \times 100$] ont été calculés pour chaque patient. Avec le questionnaire Oswestry, un seuil d'amélioration de 50 % a déjà été utilisé (Hicks *et al.*, 2005) et plus récemment justifié comme un critère valable pour définir un succès clinique

chez les patients souffrant de lombalgie (Fritz *et al.*, 2009). Un changement cliniquement important de 10 points dans les scores OSW a aussi été considéré (Ostelo *et al.*, 2008). En utilisant ces critères, les trois sous-groupes ont été définis comme suit :

- Succès : $\Delta\text{OSW} \% \geq 50 \%$
- Amélioration cliniquement significative : $\Delta\text{OSW} \% < 50 \%$, mais $\Delta\text{OSW} \geq 10$
- Échec : $\Delta\text{OSW} \% < 50 \%$ et $\Delta\text{OSW} < 10$

Des modèles de régression logistique ont ensuite été développés afin de définir, sur la base des mesures initiales (PHT et la plupart des mesures PSY à T0), les caractéristiques des patients qui ont eu un succès thérapeutique à T8. Il convient de mentionner ici qu'à la suite d'explorations dans le but d'exploiter toute l'information disponible dans nos variables continues, il a été choisi de ne pas limiter les investigations aux méthodes conventionnelles de recherche des points de coupure (p. ex. : Courbe ROC, aire sous la courbe, Index de Youden), mais de les élargir à d'autres méthodes moins conventionnelles basées sur la théorie de la décision statistique et sur la théorie de l'information. Il s'est avéré que les méthodes conventionnelles basées sur un critère statistique libéral ($P < 0,20$) ont écarté une proportion assez importante de prédicteurs. Or, selon les nouvelles méthodes, plusieurs d'entre eux possédaient une valeur informative modérée à haute. La valeur ajoutée de ces méthodes semble être liée d'une part au contenu potentiellement informatif des tests diagnostiques utilisés, et d'autre part à la prise en compte des divergences des distributions de probabilité. Ces deux concepts méthodologiques récemment introduits dans le domaine diagnostic (Benish, 2020), mais ignorés dans les méthodes classiques, rendent la recherche et la sélection des seuils de décision clinique beaucoup plus précises, efficaces et novatrices. Pour des fins de concision, l'exposition de ces processus complexes ne sera pas rapportée dans le présent rapport, mais sera abordée dans une prochaine publication scientifique.

Les étapes de ces processus sont décrites sommairement comme suit :

1. Six méthodes ont été utilisées afin d'identifier les points de coupure les plus informatifs de nos variables continues et pouvant posséder un certain potentiel de discrimination entre les groupes succès et échec préalablement définis. La méthode 1 (ROC) est déduite de la courbe ROC (Receiver Operating characteristic) qui est en effet le tracé représentant le taux des vrais positifs au taux des faux positifs. Cette méthode considère le point de la courbe situé à l'extrémité gauche de la courbe. La méthode 2 (Md) rapporte une simple coupure de la série à sa valeur médiane. La méthode 3 (IY) utilise la valeur maximale de l'index de Youden (1950) comme point informatif; cette méthode rapporte généralement (mais pas toujours) la même information que la courbe ROC. La méthode 4 ($\Delta\text{MinSnSp}$) se rapporte à la différence minimale entre le taux des vrais positifs et le taux des vrais négatifs, permettant ainsi une réduction maximale des faux diagnostics (faux positifs et faux négatifs). La méthode 5 (LRmax) est basée sur le rapport de la vraisemblance positive ou LR+ (positive likelihood ratio), qui est le rapport du taux de vrais positifs au taux de faux positifs. Une valeur élevée du LR+ indiquerait dans quelle mesure elle est importante dans la modification de la

probabilité générée par un test de diagnostic positif. La méthode 6 (KLf/g ou KLg/f) est basée sur les distances de probabilités (Dfllg) et (Dgllf) de Kullback-Leibler (Lee, W. C., 1999).

2. L'utilisation et le croisement de ces différentes méthodes nous ont permis un meilleur sondage de nos variables et de la sélection des points de coupure les plus informatifs. Des investigations conjointes ont été également menées pour examiner les points de coupure sélectionnés en termes clinique et physiologique, et leur concordance avec la littérature disponible, le cas échéant. Les caractéristiques statistiques et fonctionnelles des variables sélectionnées ont aussi été examinées à la lumière des informations apportées par la sensibilité, la spécificité et les rapports de vraisemblances, de même que leurs capacités de discrimination exprimées par l'aire sous la courbe ROC et la négociation entre les taux de faux diagnostics.
3. Pour chaque variable dichotomique, une analyse univariée (khi carré) entre les sous-groupes « succès » et « échec » a été réalisée pour sélectionner les variables qui ont un certain potentiel de prédiction, en utilisant un critère statistique libéral ($P < 0,20$). Le rapport de vraisemblance positif (LR+) correspondant a aussi été considéré de sorte qu'une variable présentant un $LR+ \geq 2$ était aussi acceptée, même si le Khi carré donnait un $P \geq 0,20$, ce qui s'est produit dans de rares cas. De toute manière, seuls les modèles multivariés qui contiennent des prédicteurs tous statistiquement significatif ont été conservés.
4. Une modélisation hiérarchique en trois étapes a été effectuée. À chacune de ces étapes, une méthode pas-à-pas descendante (backward stepwise approach) a été utilisée. Ces trois étapes sont décrites comme suit :
 - a. Étape 1 : considération des variables de classe A, soient les variables qui doivent être considérées pour l'ajustement des modèles (variables confondantes : âge, sexe, indice de masse corporelle ou IMC), la génétique étant à la source de différences liées à l'âge et au sexe (Freidin *et al.*, 2021), ainsi que les variables qui sont spécifiquement (théoriquement) associées à ce traitement, soient en relation directe ou indirecte avec la stabilité lombaire;
 - b. Étape 2 : en plus des variables de classe A, considération des variables qui peuvent influencer les résultats grâce à l'adhésion au programme d'exercice à domicile (variables de classe B - potentiellement liés à l'adhésion);
 - c. Étape 3 : en plus des variables des classes A et B, considération des variables qui peuvent être associées à n'importe quel programme d'exercice (variables de classe C – non spécifiques au traitement).
5. Ensuite, différentes statistiques permettant de juger de la performance des modèles multivariés retenus, dont les coefficients de détermination (R^2 de Nagelkerke), les scores de sensibilité et de spécificité ainsi que les valeurs prédictives et les rapports de vraisemblance positifs et négatifs ont été calculés.

Les variables des classes A, B et C sont décrites dans les **annexes A.1 et A.2**.

L'interprétation des résultats à chaque étape aidera à définir la signification clinique de la RPC. Étant donné que le but de la RPC consiste à prédire si la condition d'un patient s'améliorera ou pas avec ce programme d'exercice, les facteurs liés à l'adhésion au programme à domicile, considérés à l'étape 2, risquent d'être importants. En effet, si ces

facteurs sont problématiques, il faudra d'abord y remédier avant de débiter le programme d'exercice. Advenant que des variables de classe C aient aussi une valeur prédictive additionnelle (étape 3), il faudra conclure qu'elles sont associées à tout programme d'exercice.

3.5.2 Objectif 2 – Étude des mécanismes d'action (hypothèses 2 et 3)

Hypothèse 2 (pour variables PHT, PSY et NRM) : Les sous-groupes succès (GS) et échec (GE) auront été définis précédemment chez les patients, de même que le sous-groupe de sujets sains ou contrôles (GC) pour les variables NRM. Pour les mesures PHT, PSY et NRM de nature continue, des ANOVA à deux voies pour mesures répétées sur le facteur TEMPS ont été modulées en fonction des sous-groupes et temps de mesure disponibles :

- Pour les mesures PHT : 2 SOUS-GROUPES (GE: $n = 45$; GS : $n = 54$) $\times$ 2 TEMPS (aux temps T0 et T8) ;
- Pour les mesures PSY : 2 SOUS-GROUPES (GE: $n = 45$; GS : $n = 54$) $\times$ 4 TEMPS (aux temps T0, T4, T8 et T34) ;
- Pour les mesures NRM : 3 SOUS-GROUPES (GC : $n = 30$; GE: $n = 31$; GS : $n = 40$) $\times$ 2 TEMPS (aux temps T0 et T8). Ici, étant donné qu'il y avait un groupe contrôle (sujets sains) n'ayant pas eu le traitement, il n'était pas possible d'attribuer l'effet TEMPS au temps ou au traitement. Par contre, l'étude de l'interaction STATUT $\times$ TEMPS permettait d'exclure l'apprentissage des tests comme explication possible de l'effet TEMPS chez les patients.

Pour ce qui est des variables PHT et PSY de nature catégorique, un tel modèle statistique mixte n'est pas disponible alors seulement le facteur TEMPS a été étudié, pour chaque sous-groupe GE et GS séparément. Le test de McNemar a ainsi été appliqué aux variables PHT (temps T0 par rapport à T8) alors que le test Q de Cochran fût appliqué aux variables PSY pour comparer les quatre temps de mesure. Étant donné qu'il n'y avait pas de groupe contrôle composé de patients n'ayant pas eu le traitement, il était impossible d'attribuer l'effet TEMPS au temps ou au traitement, c'est la raison pour laquelle il a été choisi d'écrire « effet temps/traitement » dans le texte qui suit portant sur les variables PHT et PSY.

De manière à permettre une interprétation plus claire de la grandeur des effets détectés, la taille d'effet a été calculée à l'aide de formulations analogues à la distance d de Cohen, soit les formulations g_s (pour groupes indépendants) et g_{av} (pour mesures répétées) de Hedges, permettant de comparer les effets entre des effets intra- et intersujets (Lakens, 2013). Tout comme le d de Cohen, une valeur g aux alentours de 0,2 est interprétée comme un effet « faible », 0,5 « moyen » et 0,8 « fort ». Afin de faciliter l'interprétation, les valeurs g pour décrire l'effet TEMPS ont été calculées de manière à ce que les valeurs négatives indiquent une diminution dans le temps.

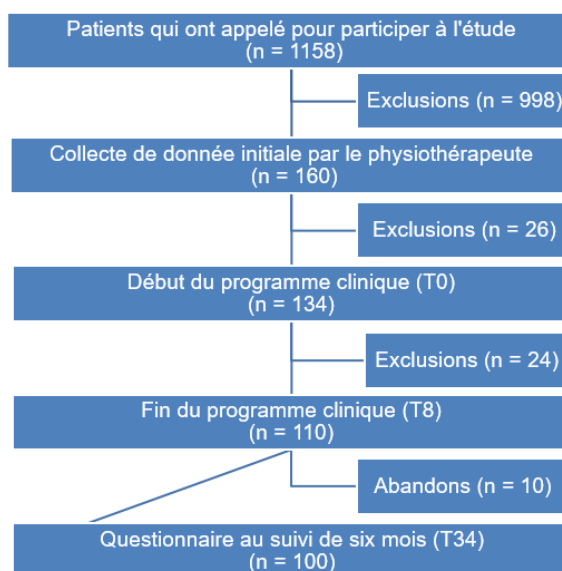
Hypothèse 3 (prédiction de l'adhésion avec les mesures PSY). Afin de connaître les variables qui ont le potentiel de prédire l'adhésion, soit celles qui pourraient être utilisées à l'étape 2 des régressions hiérarchiques décrites dans la section 2.5.1 précédente (dérivation des RPC de succès et d'échec), un modèle multivarié (procédure PROC MIXED, SAS 9.3; SAS Institute, Cary NC) a été créé pour lier ces variables à l'adhésion. Ce modèle a la particularité de tenir compte des mesures répétées faites dans le temps, c.-à-d. les variables indépendantes mesurées à T0, T4, T8 et T34, la variable dépendante (adhésion) mesurée à T8 et T34 et la différence (T8-T34) entre la variable dépendante mesurée à T8 et T34.

4. RÉSULTATS

4.1 Recrutement des participants

Seulement six patients ont été recrutés dans les cliniques, car ceux-ci débutent généralement un programme d'exercice avant la fin de la phase aiguë. Les patients ont été recrutés principalement par le biais d'annonces dans les journaux, comme détaillé avec les statistiques suivantes. Ainsi, 1158 appels (1^{re} phase : 714 ; 2^e phase : 444) ont été reçus, dont 160 (77 + 83) ont permis de recruter des participants pour la première évaluation clinique (**figure 4**). Puis, 26 (13+13) d'entre eux ont été exclus parce qu'ils ne répondaient pas à certains critères (12 % OSW ou signes neurologiques). Parmi les 134 (64 + 72) patients qui ont débuté le programme d'exercice, 110 (48 + 62) l'ont complété et ont été réévalués (questionnaires et examen physique post-traitement), toutes ces mesures étant nécessaires pour répondre à l'objectif 1 (dérivation préliminaire des RPC). Cette attrition (n = 24 patients ; soit 18 %) s'explique le plus souvent par des facteurs dits « non évitables » (n = 16) par l'équipe de recherche (début grossesse, déménagement, chirurgie, décès dans la famille, apparition de douleurs cervicales, plus de réponse après le début de la pandémie COVID-19), mais aussi pour des raisons évitables (7 pour manque d'assiduité aux rendez-vous; 1 parce qu'il ne percevait pas assez d'amélioration). Au suivi à 6 mois, une attrition additionnelle de 10 patients a été observée.

Figure 4. Diagramme de suivi des participants, de la phase de recrutement jusqu'à la fin de l'étude.



Les caractéristiques démographiques, anthropométriques et cliniques des patients étaient équivalentes entre les hommes (n = 50) et les femmes (n = 60), à l'exception de la taille et

la masse, comme attendu (**tableau 3**). La dramatisation de la douleur était légèrement plus élevée chez les hommes. Dans l'ensemble, ces patients avaient une incapacité et une douleur modérées ainsi que des caractéristiques psychologiques sous les seuils de signification clinique, comme révélé avec les échelles Dramatisation/douleur (seuil : 30/52) et IDP (seuil : 26 %), ou tout près ou sur ces seuils, tel qu'observé avec le STarT Back (seuil : 4/9).

La durée de la douleur était distribuée de la façon suivante à travers les différents intervalles proposés : (1) moins de 1 mois (n = 0); (2) 1-3 mois (n = 2); (3) 3-6 mois (n = 2); (4) 6 mois-1 an (n = 12); (5) 1-5 ans (n = 40); (6) plus de 5 ans (n = 54). L'échantillon était donc composé à 97 % (107/110) de patients avec douleurs chroniques (3 mois ou plus).

Tableau 3. Description des caractéristiques démographiques, anthropométriques et cliniques des 110 patients

Variables	Hommes (n = 50)		Femmes (n = 60)		Test-t Valeur P
	Moyenne	(Écart-type)	Moyenne	(Écart-type)	
Âge (années)	44	(11)	43	(13)	0,640
Taille (m)	1,75	(0,07)	1,64	(0,06)	< 0,001
Masse (kg)	82	(16)	71	(13)	< 0,001
IMC (kg/m ²)	26,6	(5,0)	26,4	(4,5)	0,758
Oswestry (%)	25,6	(9,0)	27,8	(10,6)	0,251
Int-Douleur (score /10)	5,2	(1,5)	5,0	(1,2)	0,450
Dramatisation/douleur (score /52)	22,9	(12,0)	18,3	(11,7)	0,045
IDPtot (score/100)	23,7	(15,0)	24,3	(13,4)	0,824
STarT Back (score/9)	4,4	(2,0)	3,7	(2,0)	0,104

IMC : Indice de masse corporelle; Oswestry : échelle Oswestry de perception d'incapacité; IDPtot : Indice de détresse psychologique – score total; STarT Back : Outil de dépistage (pronostic douleur chronique).

4.2 Effets du programme d'exercice

Durant les huit premières semaines, tous les patients ont atteint la phase 3, c.-à-d. un niveau modéré/avancé de contrôle des déficiences et fonctions, où l'accent est placé sur le développement de l'endurance. L'adhésion aux exercices à domicile, telle que mesurée à T8 et T34, était élevée avec un ratio de $0,81 \pm 0,28$ (IC 95 % : 0,76 - 0,87) et $0,51 \pm 0,40$ (IC 95 % : 0,44 - 0,59), respectivement.

Le programme a produit des effets significatifs dans le temps sur toutes les mesures de résultats primaires (**tableau 4**). Des progrès ont été observés autant au regard de la douleur, des incapacités (OSW) et des changements globaux (ChangGlob), et cela dès le mi-traitement (T4) et les gains acquis à T8 se sont maintenus au suivi de 6 mois (T34), à l'exception des changements globaux. La satisfaction des patients à l'endroit du traitement s'est avérée très élevée (92 %), telle que mesurée à T8.

Tableau 4. Effet du temps/traitement sur les mesures de résultats primaires (OUT), chez les 100 patients qui ont fait le suivi de 6 mois

Variables	Moyenne (écart-type)				ANOVA Valeur <i>P</i>	Test Post-hoc
	T0	T4	T8	T34		
Oswestry (%)	27 (9)	19 (11)	15 (11)	14 (11)	< 0,001	T4, T8, T34
Int-Douleur (/10)	5,0 (1,3)	3,5 (1,3)	2,7 (1,4)	2,9 (1,8)	< 0,001	T4, T8, T34
ChangGlob (-3 à 3)		1,4 (0,9)	2,0 (1,0)	1,6 (1,3)	< 0,001	T8 > T4, T34
Satisfaction-PT (%)			92 (12)			

Sous-groupes à T8 : Les patients correspondant aux critères de succès, d'amélioration cliniquement significative et d'échec étaient respectivement au nombre de 54 (23 H + 31 F), 11 (4 H + 7 F) et 45 (23 H + 22 F). Les mesures OUT obtenues pour ces sous-groupes de patients sont présentées au **tableau 5**. Les *d* de Cohen correspondant à la diminution de OSW et Int-Douleur à T8 sur l'ensemble du groupe (*n* = 110) étaient respectivement de -1,24 et -1,70. Ces effets étaient plus élevés pour le sous-groupe succès (-2,53 et -2,44, respectivement) que pour les sous-groupes amélioration clinique (-1,86 et -1,56) et échec (-0,21 et -1,36).

Tableau 5. Mesures de résultats primaires (OUT) chez les 110 patients qui ont complété le programme d'exercice de huit semaines, classés selon le niveau de succès du traitement

Variables	Temps	Tous (<i>n</i> = 110)		Échec (<i>n</i> = 45)		Amélioration (<i>n</i> = 11)		Succès (<i>n</i> = 54)	
Oswestry (%)*	T0	27	(10)	24	(9)	36	(9)	27	(10)
	T8	14	(11)	22	(10)	21	(7)	7	(5)
Int-Douleur (/10)	T0	5,0	(1,3)	5,3	(1,4)	5,2	(1,6)	4,8	(1,2)
	T8	2,7	(1,4)	3,4	(1,4)	3,0	(1,2)	2,1	(1,0)
ChangGlob (-3 à 3)									
	T8	2,1	(0,9)	1,6	(1,0)	2,4	(0,8)	2,4	(0,7)
Satisfaction-PT (%)									
	T8	92	(12)	92	(8)	95	(4)	92	(15)

* Variable ayant servi de critère pour déterminer l'appartenance aux sous-groupes « Échec », « Amélioration », « Succès ».

Sous-groupes à T34 : Les patients correspondant aux critères de succès, d'amélioration cliniquement significative et d'échec étaient respectivement au nombre de 53 (25 H + 28 F), 11 (1 H + 10 F) et 36 (16 H + 20 F). Les mesures OUT obtenues pour ces sous-groupes de patients sont présentées au **tableau 6**, à l'exception de la satisfaction du patient, car il n'y avait plus d'intervention de prodiguée à ce moment. Les *d* de Cohen correspondant à la diminution de OSW et Int-Douleur à T34 sur l'ensemble du groupe (*n* = 100) étaient respectivement de -1,24 et -1,32. Ces effets étaient plus élevés pour le sous-groupe succès (-3,01 et -2,43, respectivement) que pour les sous-groupes amélioration clinique (-1,37 et -1,22) et échec (0,0 et -0,72).

Tableau 6. Mesures de résultats primaires (OUT) chez les 100 patients qui ont complété les mesures prises à six mois de suivi (T34), classés selon le niveau de succès du traitement

Variables	Temps	Tous (n = 100)		Échec (n = 36)		Amélioration (n = 11)		Succès (n = 53)	
Oswestry (%)*	T0	27	(10)	23	(8)	37	(9)	27	(9)
	T34	14	(11)	23	(10)	24	(10)	6	(4)
Int-Douleur (/10)	T0	5,0	(1,3)	5,1	(1,5)	5,7	(1,1)	4,9	(1,2)
	T34	3,0	(1,7)	3,9	(1,8)	3,8	(1,9)	2,1	(1,1)
ChangGlob (-3 à 3)									
	T34	1,6	(1,3)	0,8	(1,3)	1,5	(1,1)	2,1	(0,9)

* Variable ayant servi de critère pour déterminer l'appartenance aux sous-groupes « Échec », « Amélioration », « Succès ».

4.3 Objectif 1 – Dérivation de RPC de succès

4.3.1 Sélection des variables prédictives potentielles

Les variables continues ont été dichotomisées à l'aide des différentes procédures décrites dans la section statistique, menant parfois à différents points de coupure pour une même variable. Seules variables qui ont démontré une tendance ($P < 0,20$) à discriminer les groupes succès et échec (test Khi carré) ou ayant un $LR+ > 2$ ont été retenues comme candidates. Ces variables sont présentées à l'**annexe B**. Les variables candidates à la RPC de succès à T8 sont présentées dans le **tableau B-1** pour les variables initialement de nature continue et dans le **tableau B-2** pour les variables initialement de nature dichotomique. Au total, 22 variables de classe A, 9 variables de classe B et 10 variables de classe C étaient candidates pour passer à l'étape des analyses multivariées. Les variables candidates à la RPC de succès à T34 sont présentées dans les **tableaux B-3** (variables initialement continues) et B-4 (variables initialement dichotomiques), pour un total de 17, 8 et 7 variables pour les classes A, B et C respectivement.

4.3.2 Analyses multivariées

4.3.2.1 Règle de prédiction clinique du succès à la fin du traitement (T8)

Neuf modèles (régressions logistiques) de succès ont été développés. Les quatre modèles les plus intéressants (modèles 3, 5, 8 et 9), c.-à-d. (1) ne comportant pas de variables qui ont presque atteint le niveau de signification statistique et (2) qui atteignent un $LR+$ supérieur à la RPC de succès de l'étude préliminaire ($LR+ = 4,0$) de Hicks *et al.* (2005), ont été retenus. Les modèles 1 et 2, bien que plus simples d'utilisation (deux ou trois variables), ne permettaient d'atteindre que des $LR+$ de 4,17 et 4,44, respectivement. Les variables (ou prédicteurs) des quatre modèles sélectionnés, les coefficients bêta qui leur sont associés ainsi que leur rang en termes de contribution relative à chaque modèle, sont présentés dans le **tableau 7**. En somme, les deux premiers modèles (modèles 3 et

5) ne sont constitués que de variables de classe A alors que les deux derniers (modèles 8 et 9) comportent également des variables de classe B. Donc aucun modèle n'inclut des variables de classe C.

Les indicateurs de performance diagnostique et d'ajustement des quatre modèles sélectionnés sont présentés dans le **tableau 8**. Rappelons que LR+ est l'indicateur statistique le plus important pour une RPC de succès. Les modèles retenus, bien que plus complexes (quatre à sept variables offrent des performances supérieures (LR+ entre 7,91 et 21,7) à la RPC de succès de l'étude préliminaire (LR+ = 4,0) de Hicks *et al.* (2005).

Tableau 7. Les coefficients des modèles prédictifs de succès à T8 et leurs intervalles de confiance à 95% (IC 95%).

Variables sélectionnées	Modèle 3 (4 variables)	Modèle 5 (5 variables)	Modèle 8 (6 variables)	Modèle 9 (7 variables)
	B (IC 95 %)	B (IC 95 %)	B (IC 95 %)	B (IC 95 %)
TPP-portée $\geq 0,75/1$	1,78 (0,32 - 3,23) ³	1,63 (0,095-3,16) ³	1,86 (0,24 - 3,48) ⁵	1,99 (0,30 - 3,67) ⁵
HanRot-Pas-Max positif	1,58 (0,59 - 2,57) ³	1,79 (0,76-2,82) ⁴	2,02 (0,92 - 3,12) ⁶	2,15 (1,00 - 3,30) ⁶
Mvt-Abe positif	1,40 (0,30 - 2,50) ²	1,60 (0,45 - 2,76) ²	1,55 (0,34 - 2,75) ⁴	1,50 (0,28 - 2,72) ⁴
Beighton < 5/9	1,65 (0,00 - 3,30) ¹	1,98 (0,45 - 2,76) ¹	2,20 (0,37 - 4,03) ¹	2,45 (0,56 - 4,34) ²
TEM-dorsaux $\geq 225,56$ s		2,33 (0,29 - 3,67) ¹	2,33 (0,10 - 4,55) ²	2,21 (0,95 - 4,47) ⁴
PercMaladie $\leq 38,5$ /80			1,49 (0,19 - 2,78) ³	1,85 (0,44 - 3,26) ³
FABQ- AP $\geq 6,5/24$				2,03 (0,13 - 3,93) ¹
Constante	-2,48	-3,05	-3,64	-5,81

B : Coefficient Beta (Log odds); ()ⁿ Classement des prédicteurs selon la statistique Adéquation (Adequacy) (n = 1 signifie qu'il représente le prédicteur le plus influent; le même rang peut être attribué lorsque cet indicateur donne le même score). La statistique adéquation donne la valeur explicative individuelle du prédicteur et donc sa force explicative. Il s'agit du ratio -2log - Vraisemblance (-2log - likelihood ou -2LL) du prédicteur par rapport à - 2LL du modèle complet, et représente la proportion de la log-Vraisemblance du modèle complet (considérant tous les prédicteurs) qui est expliquée par chaque prédicteur individuellement.

Un autre indicateur apporte une information complémentaire importante, soit la probabilité de succès en post-traitement estimée avec la valeur prédictive positive (VP+). Cet indicateur est analogue à ce qui est calculé en utilisant le nomogramme de Fagan, comme fait dans l'étude de référence de Hicks *et al.* (2005). Pour des fins de comparaison, l'estimation donnée par le nomogramme de Fagan (sera nommée VP-Fagan) a également été rapportée dans le **tableau 8**, car elle est communément utilisée dans la pratique clinique. VP-Fagan permet d'estimer que l'utilisation des différentes RPC de succès (modèle 3 : LR+ = 7,91; modèle 5 : LR+ = 8,33; modèle 8 : LR+ = 11,25; modèle 9 : LR+ = 21,7) permettrait de passer de 49 % de succès (sans RPC) à un succès post-traitement bien supérieur, soit 88 % (IC 95 % : 68-96), 89 % (IC 95 % : 69-97), 91 % (IC 95 % : 76-97) et 95 % (IC 95 % : 78-99), respectivement (notez que les VP+ sont légèrement supérieures de 1 à 2 %). Par contre, pour passer d'une probabilité en post-traitement de 88 % (modèle 3) à 95 % (modèle 9), il faut ajouter trois variables à la RPC (TEM-dorsaux, PercMaladie et Peurs/croyances-AP), ce qui allonge significativement l'évaluation clinique pour un gain relativement mineur (de 7 %) en termes de probabilité de succès en post-traitement. Finalement, il est important de noter que la spécificité du modèle 3 (95,6 %) pour le modèle ayant un LR+ de 7,91 est presque la même que dans le modèle 9 (spécificité = 97,8 %) ayant un LR+ de 21,7. Conséquemment, c'est le modèle 3 du **tableau 7**, comprenant 4 variables de classe A, qui est retenu pour en faire la RPC de succès. À noter que les autres indicateurs statistiques lui sont aussi favorables. Les quatre tests à réaliser pour cette RPC de succès sont illustrés et détaillés à la **figure 5**.

Tableau 8. Indicateurs de performance diagnostique et d'ajustement des modèles prédictifs à T8 (IC 95 %).

	Modèle 3 (4 variables)	Modèle 5 (5 variables)	Modèle 8 (6 variables)	Modèle 9 (7 variables)
Indicateurs de performance diagnostique				
• Sensibilité	35,2 (22,4 - 47,9)	37,0 (24,2 - 49,9)	50,0 (36,7 - 63,3)	48,1 (34,8 - 61,5)
• Spécificité	95,6 (89,5 - 100)	95,6 (89,5 - 100)	95,6 (89,5 - 100)	97,8 (93,5 - 100)
• Rapport de vraisemblance positif (LR+)	7,91 (1,95 - 32,2)	8,33 (2,06 - 33,7)	11,25 (2,83 - 44,8)	21,7 (3,1 - 153,5)
• Rapport de vraisemblance négatif (LR-)	0,68 (0,55 - 0,83)	0,66 (0,53 - 0,82)	0,52 (0,40 - 0,69)	0,53 (0,41 - 0,69)
• Valeur prédictive positive (VP+)	90,5 (77,9 - 100)	91,0 (78,9 - 100)	93,1 (83,9 - 100)	96,3 (89,2 - 100)
• Valeur prédictive négative (VP-)	55,1 (44,1 - 66,2)	55,8 (44,7 - 66,9)	61,4 (50,0 - 72,8)	61,1 (49,8 - 72,4)
• Aire sous la courbe (AUC), valeur <i>P</i> *	0,78 (0,68 - 0,87)	0,79 (0,71 - 0,88)	0,83 (0,75 - 0,91)	0,85 (0,78 - 0,93)
• Probabilité succès post-traitement (%) †	88 (68-96)	89 (69-97)	91 (76-97)	95 (78-99)
Indicateurs explicatifs et d'ajustement				
• Coefficient R ² de Nagelkerke (%)	32,2	37,8	42,9	47,5
• Test de Hosmer et Lemeshow (χ^2 , <i>p</i>)	($\chi^2=4,31$; <i>P</i> = 0,37)	($\chi^2=3,11$; <i>P</i> = 0,79)	($\chi^2=2,31$; <i>P</i> = 0,89)	($\chi^2= 2,98$; <i>P</i> = 0,89)

* Les valeurs *P* sont toutes < 0,001; Test de Hosmer et Lemeshow : test de la qualité d'ajustement (doit être non significatif).

† Estimée avec le nomogramme de Fagan (<http://araw.mede.uic.edu/cgi-bin/testcalc.pl>), connaissant LR+, LR-, et en estimant la prévalence (ici 54/110 patients, soit 49 %). Pour le modèle 3, ceci indique que la probabilité de succès passe de 49 % (sans utilisation de la RPC) à 88 % (avec utilisation de la RPC).

Pour la RPC de succès sélectionnée (modèle 3 comprenant 4 variables de classe A), le nombre de participants des groupes succès et échec a été déterminé en fonction du nombre de tests positifs de la RPC de succès (**tableau 9**). À l'aide des effectifs (nombre de patients) du **tableau 9**, les indicateurs de performance diagnostique correspondants ont également été produits (**tableau 10**). L'information contenue dans ces deux tableaux permet de conclure que la meilleure combinaison pour prédire le succès du traitement apparaît lorsque deux prédicteurs (ou tests cliniques) ou plus sont positifs (ou atteints), car cette condition génère un LR+ cliniquement plus important, soit 17,9, ainsi que la plus haute probabilité de succès, soit 95,6 %. Bien que la considération de trois prédicteurs ou plus semble mener à de meilleurs résultats (LR+ de 23,9), l'intervalle de confiance qui lui est associé est très grand (1,6 à 365,8), soulignant un manque de précision. De toute manière, la probabilité de succès est à peine supérieure (96,7 %) à la solution à deux prédicteurs ou plus (95,6 %). Comparativement au modèle global (LR+ de 7,9), il est normal que le nombre croissant de tests cliniques, même générant des LR+ importants, s'accompagne par une réduction marquée des effectifs (**tableau 9**), ce qui réduira, comme rapportée ici, la précision des estimations.

Tableau 9. Nombre de participants des groupes succès et échec en fonction du nombre de tests positifs de la RPC de succès à T8 (modèle 3 avec 4 variables de classe A)

Nombre de tests positifs	Nombre de participants du groupe succès (n = 54)	Nombre de participants du groupe échec (n = 45)
1 ou plus	54	40
2 ou plus	43	15
3 ou plus	14	2
4	1	0

Tableau 10. Indicateurs de performance diagnostique (intervalles de confiance à 95 %) associés au nombre de prédicteurs présents dans la RPC de succès à T8 (modèle 3 à 4 variables de classe A)

Nombre de prédicteurs présents	Sensibilité	Spécificité	Rapport de vraisemblance positif (LR+)	Probabilité de succès (%) ou VPP après le PESL
1 ou plus	18,5 (8,2 - 28,9)	97,8 (93,5 - 100)	8,3 (1,1 - 62,6)	90,9 (73,9 - 100)
2 ou plus	79,6 (68,9 - 90,4)	95,5 (89,5 - 100)	17,9 (4,6 - 69,9)	95,6 (89,5 - 100)
3 ou plus*	85,3 (68,5 - 100)	96,6 (86,8 - 100)	23,9 (1,6 - 365,8)	96,7 (87,7 - 100)
4	-	-	-	-

PESL : Programme d'exercices de stabilisation lombaire.

* Avec trois prédicteurs ou plus, les effectifs sont considérablement réduits (14 succès et 2 échecs), comme décrit dans le **tableau 9**, rendant le calcul des indicateurs statistiques impossible (groupe échec = 2). Comme fait dans d'autres études (Flynn *et al.*, 2002; Stolze *et al.*, 2012), une approximation des indicateurs a donc été estimée, à titre indicatif, en utilisant les effectifs du groupe échec correspondant à la condition avec deux prédicteurs ou plus (n = 15).

Figure 5. Tests PHT retenus dans la RPC de succès à la fin du traitement (au temps T8).

		
TPP-portée (Novy <i>et al.</i>, 2002): Un ruban est collé horizontalement sur le mur à la hauteur des épaules. Position initiale du patient : alignement de l'acromion avec le 0 du ruban. La longueur du bras est alors mesurée en demandant d'allonger les bras avec les poings fermés (sans sac) et en prenant la mesure de la position des jointures métacarpo-phalangiennes. Ensuite, le patient doit tenir un poids (5 % du poids du patient jusqu'à concurrence de 5 kg) près du corps à la hauteur des épaules pour ensuite allonger les bras à l'horizontale vers l'avant jusqu'à la limite de la tolérance du patient. La distance maximale atteinte est mesurée et divisée par la longueur du bras précédemment mesurée, pour donner un ratio. Le test est positif si ce ratio est $\geq 0,75$.	HanRot-Pas-Max (Van Dillen <i>et al.</i>, 2003): Abduction et rotation latérale de la hanche, de façon passive (générée par le thérapeute). Les symptômes produits sont comparés aux symptômes en position couchée sur le dos (augmentation [1], diminution [-1], stable [0]). Le test est fait pour chaque membre inférieur et il est positif si au moins un des deux côtés produit une augmentation des symptômes.	Mouvements aberrants (Hicks <i>et al.</i>, 2003): En position debout, le patient fléchit le tronc vers l'avant autant que possible. Un seul essai est autorisé. Si une des 5 anomalies du mouvement est présente, le test est positif : - Arc douloureux en flexion : symptômes ressentis pendant le mouvement à un point particulier du mouvement (pas présents avant ou après ce point); - Arc douloureux au retour : symptômes n'apparaissent que lors du retour à la position verticale; - Signe de Gower (appui sur les cuisses) : pousse sur les cuisses ou une autre surface avec les mains pour aider au retour en position verticale; - Ressaut d'instabilité (<i>instability catch</i>) : toute accélération ou décélération soudaine du mouvement du tronc ou tout mouvement se produisant en dehors du plan de mouvement principal (par exemple, flexion latérale ou rotation); - Inversion du rythme lombo-pelvien : Le patient utilise une flexion des genoux pour ramener le bassin à la verticale avant de se redresser.

Figure 5 (suite). Tests PHT retenus dans la RPC de succès à la fin du traitement (au temps T8).



Laxité ligamentaire généralisée : Échelle de Beighton, comme décrit par Hicks *et al.* (2003):

Ces tests sont généralement réalisés avec une évaluation visuelle seulement alors lorsque ce n'est pas clair, le test est négatif. Comme illustré dans certaines images, il arrivait parfois qu'un goniomètre fût utilisé lorsqu'un doute subsistait avec l'évaluation visuelle. Quatre tests sont évalués à droite et à gauche et un point est attribué pour chaque test que le sujet peut effectuer. Les tests bilatéraux sont :

- Hyperextension passive du coude supérieure à 10°.
- Hyperextension passive du cinquième doigt à plus de 90°.
- Abduction passive du pouce pour entrer en contact avec l'avant-bras.
- Hyperextension passive des genoux supérieure à 10°.

Le test final est la capacité à fléchir le tronc, avec les pieds à la largeur des hanches, et à placer les deux mains à plat sur le sol sans fléchir les genoux. Les scores possibles sont de 0 à 9, le plus élevé indiquant une plus grande laxité.

Le test est positif si le score est < 5/9.

4.3.2.2 Règle de prédiction clinique du succès au suivi de 6 mois (T34)

Seize modèles (régressions logistiques) de succès ont été développés. Les six modèles les plus intéressants (critères énoncés dans la section précédente) sont décrits dans le **tableau 11** (variables et coefficients bêta) et le **tableau 12** (indicateurs de performance diagnostique et d'ajustement). En somme, les deux premiers modèles (modèles 2 et 3) ne sont constitués que de variables de classe A, les deux suivants (modèles 8 et 12) de variables de classe A et B alors que les deux derniers (modèles 14 et 16) comportant aussi des variables de classe C.

Sur la base des LR+, il y a un écart important, à l'avantage des deux derniers **modèles (14 et 16)**, mais ces modèles comportent des variables de classe C qui n'ont pas de relation théorique avec l'objet de cette étude (instabilité lombaire). Ces modèles sont éliminés et seront utilisés à titre d'éléments de discussion. À noter que la VP+ n'est que de 2-3 % supérieure pour ces modèles (95,7 et 95,8 %) que pour les modèles 3 et 12 (92,3 et 93,8 %), ce qui représente un avantage négligeable du côté de la valeur prédictive. À noter également que la spécificité des 6 modèles est équivalente (97 %).

Parmi les quatre premiers modèles et toujours sur la base des LR+, ce sont initialement le modèle 3 (LR+ = 8,15) à 3 variables de classe A et le modèle 12 (LR+ = 10,2) à 6 variables (classes A et B) qui pour des raisons pratiques et cliniques ont retenu notre attention pour des analyses plus poussées. Par contre, la VP+ de ces deux modèles étant similaire (92,3 et 93,8 % respectivement), il apparaît assez clairement que le modèle le plus simple (modèle 3 à trois variables) serait à privilégier pour des raisons pratiques (temps d'évaluation par les intervenants). Le modèle 3 avait initialement été retenu pour la suite des analyses, mais ces dernières se sont avérées décevantes en termes de valeur prédictive (p. ex. : LR+ de 1,9). Le modèle 2 a finalement été retenu, n'impliquant qu'une variable de plus (âge < 43 ans), laquelle n'ayant pas véritablement d'impact sur la durée de l'évaluation clinique, mais permettant d'ajuster sur un effet confondant pouvant être lié à l'âge des patients. Les trois tests cliniques à réaliser pour cette RPC de succès sont illustrés et détaillés à la **figure 6**.

Pour la RPC de succès sélectionnée (modèle 2 comprenant 4 variables de classe A), le nombre de participants des groupes succès et échec a été déterminé en fonction du nombre de tests positifs de la RPC de succès (**tableau 13**), de même que les indicateurs de performance diagnostique correspondants (**tableau 14**). Des estimations ont dû être réalisées pour compenser les effectifs réduits rencontrés lorsque 2 ou 3 prédicteurs ou plus sont positifs, comme décrit au bas du **tableau 14**. La meilleure combinaison pour prédire le succès du traitement apparaît lorsque trois prédicteurs (ou tests cliniques) ou plus sont positifs (ou atteints), ce qui mène à un LR+ de 17,0 ainsi qu'une probabilité de succès de 85,0 %.

Tableau 11. Les coefficients des modèles prédictifs de succès à T34 et leurs intervalles de confiance à 95 % (IC 95 %)

Variables sélectionnées	Modèle 2	Modèle 3	Modèle 8	Modèle 12
	B (IC 95 %)	B (IC 95 %)	B (IC 95 %)	B (IC 95 %)
HanExt-Act-Max négatif	1,05 (0,003-2,09) ²	1,13 (0,11-2,15) ¹	1,17 (0,055-2,28) ¹	1,22 (0,02-2,42) ²
TEM-abdominaux ≥ 72,62 s	1,58 (0,38-2,77) ⁴	1,31 (0,14-2,48) ²	1,66 (0,39-2,92) ³	1,75 (0,44-3,05) ⁵
TPP-flexions ≥ 9,91 s	1,54 (0,33-2,75) ³		1,74 (0,44-3,03) ²	2,47 (0,97-3,97) ³
Âge < 43 ans	0,60 (-0,38-1,58) ¹			0,92 (-0,19-2,03) ¹
TPP-portée ≥ 0,76/1		2,07 (0,45-3,69) ³		
PercMaladie < 45 /80			1,44 (0,41-2,47) ⁴	/
PercMaladie < 44 /80			/	1,46 (0,32-2,61) ⁴
FABQ- AP < 12/24				1,44 (0,27-2,6) ⁴
EJT-Pas-Max-EDM-Moy < 93°				
EJT-Pas-Doul-EDM-Min < 68°				
Constante	-0,10	-0,58	-1,47	-2,54

Suite du tableau (2 derniers modèles)	Modèle 14	Modèle 16
	B (IC 95 %)	B (IC 95 %)
HanExt-Act-Max négatif	1,18 (0,05-2,31) ¹	1,47 (0,20-2,75) ¹
TEM-abdominaux ≥ 72,62 s	1,54 (0,23-2,85) ²	2,58 (0,96-4,20) ²
TPP-flexions ≥ 9,91 s		2,38 (0,83-3,94) ²
Âge < 43 ans		
TPP-portée ≥ 0,76/1	2,49 (0,60-4,38) ³	
PercMaladie < 45 /80	/	/
PercMaladie < 44 /80	1,07 (0,005-2,14) ²	1,68 (0,40-2,96) ³
FABQ- AP < 12/24	1,84 (0,49-3,20) ²	2,09 (0,65-3,53) ³
EJT-Pas-Max-EDM-Moy < 93°	1,80 (0,29-3,32) ¹	
EJT-Pas-Doul-EDM-Min < 68°		2,51 (0,86-4,14) ³
Constante	-3,15	-4,53

B : Coefficient Beta (Log odds); ()ⁿ Classement des prédicteurs selon la statistique Adéquation (Adequacy) (n = 1 signifie qu'il représente le prédicteur le plus influent; le même rang peut être attribué lorsque cet indicateur donne le même score). La statistique adéquation donne la valeur explicative individuelle du prédicteur et donc sa force explicative. Il s'agit du ratio -2log - Vraisemblance (-2log - likelihood ou -2LL) du prédicteur par rapport à - 2LL du modèle complet, et représente la proportion de la log-Vraisemblance du modèle complet (considérant tous les prédicteurs) qui est expliquée par chaque prédicteur individuellement.

Tableau 12. Indicateurs de performance diagnostique et d'ajustement des modèles prédictifs à T34 (IC 95 %).

	Modèle 2 (4 variables)	Modèle 3 (3 variables)	Modèle 8 (4 variables)	Modèle 12 (6 variables)
Indicateurs de performance diagnostique				
• Sensibilité	20,75 (9,84-31,67)	22,64 (11,37-33,9)	15,1 (5,5-24,7)	28,3 (16,2-40,4)
• Spécificité	97,2 (91,8-100)	97,22 (91,85-100)	97,2 (91,8-100)	97,2 (91,8-100)
• Rapport de vraisemblance positif (LR+)	7,47 (1,01-55,37)	8,15 (1,11-59,9)	5,43 (,71-41,6)	10,2 (1,41-73,8)
• Rapport de vraisemblance négatif (LR-)	0,82 (0,70-0,95)	0,80 (0,68-0,93)	0,87 (0,77-0,99)	0,74 (0,62-0,88)
• Valeur prédictive positive (VP+)	91,67 (76,0-100)	92,3 (77,8-100)	89,0 (68,4-100)	93,8 (81,9-100)
• Valeur prédictive négative (VP-)	45,45 (34,33-56,57)	46,1 (34,8-57,3)	43,8 (32,9-54,6)	47,9 (36,5-59,4)
• Aire sous la courbe (AUC), valeur <i>P</i> *	0,728 (0,62-0,84)	0,73 (0,62-0,83)	0,77 (0,67-0,87)	0,82 (0,73-0,91)
• Probabilité succès post-traitement (%) †	89 (58-98)	90 (61-98)	86 (51-97)	92 (66-99)
Indicateurs explicatifs et d'ajustement				
• Coefficient R2 de Nagelkerke (%)	22,0	23,4	30,3	39,5
• Test de Hosmer et Lemeshow (χ^2 , p)	(χ^2 = 4,15; <i>P</i> = 0,66)	(χ^2 = 0,11; <i>P</i> = 0,99)	(χ^2 = 2,87; <i>P</i> = 0,90)	(χ^2 = 1,84 ; <i>P</i> = 0,99)
	Modèle 14 (6 variables)	Modèle 16 (6 variables)		
Indicateurs de performance diagnostique				
• Sensibilité	41,5 (28,2-54,8)	43,4 (30,0-56,7)		
• Spécificité	97,2 (91,8-100)	97,2 (91,8-100)		
• Rapport de vraisemblance positif (LR+)	14,9 (2,11-106,0)	15,6 (2,21-110,6)		
• Rapport de vraisemblance négatif (LR-)	0,60 (0,48-0,76)	0,58 (0,46-0,74)		
• Valeur prédictive positive (VP+)	95,7 (87,3-100)	95,8 (87,8-100)		
• Valeur prédictive négative (VP-)	53,0 (41,0-65,1)	53,8 (41,7-66,0)		
• Aire sous la courbe (AUC), valeur <i>P</i> *	0,83 (0,75-0,92)	0,86 (0,80-0,93)		
• Probabilité succès post-traitement (%) †	94 (75-99)	95 (76-99)		
Indicateurs explicatifs et d'ajustement				
• Coefficient R2 de Nagelkerke (%)	41,6	48,7		
• Test de Hosmer et Lemeshow (χ^2 , p)	(χ^2 = 5,55; <i>P</i> = 0,70)	(χ^2 = 4,49; <i>P</i> = 0,81)		

* Les valeurs *P* sont toutes < 0,001; Test de Hosmer et Lemeshow : test de la qualité d'ajustement (doit être non significatif).

† Estimée avec le nomogramme de Fagan (<http://araw.mede.uic.edu/cgi-bin/testcalc.pl>), connaissant LR+, LR-, et en estimant la prévalence (ici 53/100 patients, soit 53 %). Pour le modèle 3, ceci indique que la probabilité de succès passe de 53 % (sans utilisation de la RPC) à 90 % (avec utilisation de la RPC).

Figure 6. Tests PHT retenus dans la RPC de succès pour le suivi de six mois (au temps T34).

		
HanExt-Act-Max (Van Dillen <i>et al.</i>, 2003): Extension active de la hanche (sans intervention du thérapeute), jusqu'à un maximum de 10° tout en gardant le genou droit. Instructions : "Levez la cuisse jusqu'à ce qu'elle soit dégagée de la table pour que je puisse passer ma main librement dessous". Les symptômes produits sont comparés aux symptômes en position couchée sur le ventre (augmentation [1], diminution [-1], stable [0]). Le test est fait pour chaque membre inférieur et il est négatif si les deux côtés ne produisent pas d'augmentation des symptômes.	TEM-abdominaux (McGill <i>et al.</i>, 1999): Le patient est en position assise, les omoplates reposant sur un support incliné à 60° du sol et le bassin étant espacé par un poing de ce support. Les genoux et hanches sont fléchis à 90°, les mains posées sur l'épaule opposée et les pieds maintenus par une sangle. Le patient doit regarder l'angle entre le plafond et le mur en face de lui durant tout le test afin de garantir un bon alignement de la colonne cervicale. On demande au patient de soulever les omoplates de la table jusqu'à ce que l'évaluateur puisse placer deux phalanges distales entre la table et la colonne thoracique. Le chronomètre est parti aussitôt. Le patient maintient la posture aussi longtemps que possible et l'évaluateur continue de rechercher un bon alignement de la colonne cervicale. On considère qu'il y a échec lorsque le patient ne peut empêcher une pression du doigt de l'évaluateur sur la colonne thoracique (arrêt du chronomètre). Le test est positif si le temps enregistré est $\geq 72,62$ s.	TPP-flexions (Novy <i>et al.</i>, 2002): En position debout, le patient fléchit le tronc vers l'avant autant que possible, comme s'il touchait ses orteils), et revient à la position debout. Ce cycle de flexion/extension doit être fait cinq fois consécutivement, aussi vite que possible. Le temps requis pour exécuter les cinq cycles est mesuré avec un chronomètre. Avant le test, le patient pratique le mouvement une ou deux fois, à un rythme lent, en guise de familiarisation et pour s'assurer d'une exécution correcte. L'évaluateur s'assure du retour complet à la position debout en plaçant son bras à la hauteur des omoplates, au début du test, et en s'assurant que le patient touche son bras (ou presque) à chaque cycle. Le test est positif si le temps enregistré est $\geq 9,91$ s. Note : Contrairement à la description de Novy <i>et al.</i> (2002) où deux essais sont réalisés, veuillez noter que le test (5 cycles) n'est effectué qu'une seule fois.

Tableau 13. Nombre de participants des groupes succès et échec en fonction du nombre de tests positifs de la RPC de succès à T34 (modèle 2 avec 4 variables de classe A).

Nombre de tests positifs	Nombre de participants du groupe succès (n = 53)	Nombre de participants du groupe échec (n = 36)
1 ou plus	52	26
2 ou plus	26	9
3 ou plus	8	1
4	0	0

Tableau 14. Indicateurs de performance diagnostique (intervalles de confiance à 95 %) associés au nombre de prédicteurs présents dans la RPC à T34 (modèle 2 à 4 variables de classe A)

Nombre de prédicteurs présents	Sensibilité	Spécificité	Rapport de vraisemblance positif (LR+)	Probabilité de succès (%) ou VP+ après le PESL
1 ou plus	84,7 (74,8-94,5)	30,8 (13,1-48,6)	1,2 (0,9-1,6)	71,0 (59,8-82,4)
2 ou plus *	98,2 (93,1-100)	64,8 (46,8-82,8)	2,8 (1,7-4,7)	73,6 (59,2-88,0)
3 ou plus †	94,4 (79,5-100)	94,4 (85,8-100)	17,0 (3,6-81,2)	85,0 (62,9-100)
4	-	-	-	-

PESL : Programme d'exercices de stabilisation lombaire.

Les symboles* et † indiquent que les effectifs du groupe échec sont considérablement réduits, comme décrit dans le **tableau 13**, rendant le calcul des indicateurs statistiques moins précis ou impossibles, selon le cas. Comme fait dans d'autres études (Flynn *et al.*, 2002; Stolze *et al.*, 2012), une approximation des indicateurs a donc été estimée, à titre indicatif, en utilisant les effectifs du groupe échec correspondant à une autre condition, co détaillé ci-dessous pour chaque cas particulier.

* Selon les prédictions, les 9 participants échec sont classés succès et donc il ne reste aucun cas échec. Le groupe échec de la condition avec 1 prédicteur ou plus (n = 26) a donc été utilisé.

† Pour des fins de puissance statistique (effectif plus important), le groupe échec de la condition avec 1 prédicteur ou plus (n = 26) a été utilisé.

4.4 Objectif 2 – Étude des mécanismes d'action

4.4.1 Évaluation physique (mesures PHT)

4.4.1.1 Comparaisons entre sous-groupes et temps de mesure (variables continues)

Les ANOVA SOUS-GROUPE (GE, GS) × TEMPS (T0, T8) ont révélé 7 interactions SOUS-GROUPE × TEMPS (**tableau 15**), dont 6 sont illustrées dans la **figure 7**. La variable non illustrée (EJT-Pas-Doul-Min), correspondant à l'angle minimal (entre les deux côtés) auquel les symptômes de douleur apparaissaient lors du test passif d'élévation de la jambe tendue, a montré une augmentation significative de cet angle seulement chez le GS (g = 0,33), ce qui s'interprète en un résultat favorable pour le GS. D'ailleurs, alors

qu'aucune différence n'était observée entre les GE ($56 \pm 13^\circ$) et GS ($56 \pm 13^\circ$) à T0, le GS ($61 \pm 13^\circ$) démontrait un angle significativement plus élevé que le GE ($55 \pm 13^\circ$) à T8 ($g = 0,41$). Dans tous les six cas illustrés (**figure 7**), les GE et GS étaient aussi équivalents à T0. À l'exception de la variable Instabilité-15Q pour laquelle aucun effet TEMPS n'a été observé pour le GS, toutes ces variables ont démontré un effet TEMPS ($P \leq 0.05$) favorable pour le GS. La nature de ces interactions faisait que cet effet favorable était plus marqué pour le GS que pour le GE, ce dernier ne démontrant pas d'effet TEMPS que pour le test d'endurance des muscles dorsaux (TEM-dorsaux).

L'effet du temps/traitement, sans interaction avec le facteur SOUS-GROUPE, a aussi été statistiquement significatif pour six autres variables (**tableau 15**), toutes indiquant une amélioration dans le temps. Bien que la taille de cet effet était petite pour la plupart d'entre-elles, soit pour TEM-abdominaux ($g = 0,15$; T0 : 56 ± 55 s; T8 : 65 ± 68 s), EJT-Pas-EDM-Min ($g = 0,18$ de T0 : $72 \pm 15^\circ$ à T8 : $74 \pm 14^\circ$), FlxBassin-EDM ($g = 0,35$ de T0 : $77 \pm 17^\circ$ à T8 : $83 \pm 15^\circ$), FlxLatLomb-EDM-Min ($g = 0,31$ de T0 : $23 \pm 9^\circ$ à T8 : $25 \pm 10^\circ$) et HanRE-Pas-EDM-Min ($g = 0,27$ de T0 : $59 \pm 10^\circ$ à T8 : $62 \pm 10^\circ$), elle était moyenne pour EJT-Act/5-Max ($g = -0,53$ de T0 : $0,91 \pm 1,20 / 5$ à T8 : $0,39 \pm 0,76 / 5$).

4.4.1.2 Comparaisons entre sous-groupes et temps de mesure (variables dichotomiques)

Contrairement aux variables continues, aucun test statistique ne permet de considérer un devis mixte pour les variables catégoriques. Bien que plusieurs tests produisent des scores dichotomiques (0 = test négatif; 1 = test positif), les scores des tests de déficience du contrôle sensorimoteur comportent trois niveaux (-1 : douleur diminuée; 0 : même douleur; 1 : douleur augmentée). Ces scores ont d'abord été dichotomisés de manière à séparer les patients qui ont un test positif (score de 1) des autres patients (0 attribué aux scores de 0 et de -1). Le test de McNemar permettant de comparer la proportion de tests positifs à T0 et à T8 a ensuite été appliqué sur l'ensemble des variables dichotomiques pour chacun des sous-groupes (GE et GS) séparément (**tableau 16**). Pour la grande majorité de ces tests, le GS a démontré une amélioration significative alors que le GE n'a démontré aucun effet (**tableau 16**).

Tableau 15. Effet du sous-groupe et du temps/traitement sur les mesures PHT (variables continues) recueillies aux temps T0 et T8

Variables*	ANOVA (Valeur <i>P</i>)		
	GROUPE (SG) (taille d'effet †)	TEMPS (T) (taille d'effet ‡)	SG × T
Variables de classe A			
Instabilité-15Q (score/75)	0,530	0,484	0,029 !!
Laxité (score/75)	0,370	0,890	0,320
Courbure (°)	0,305	0,404	0,496
TPP-portée (score/1)	0,501	0,367	0,211
TPP-assis (s)	0,801	<0,001 (g = -0,67)	0,002 !!
TPP-flexions (s)	0,512	<0,001 (g = -0,60)	0,009 !!
TPP-rouler (max 2 côtés) (s)	0,474	<0,001 (g = -0,65)	0,024 !!
TEM-lat (min 2 côtés) (s)	0,595	<0,001 (g = 0,50)	0,048 !!
TEM-abdominaux (s)	0,335	<0,001 (g = 0,15)	0,161
TEM-dorsaux (s)	0,291	<0,001 (g = 0,16)	0,010 !!
EJT-Act/5-Max	0,221	<0,001 (g = -0,53)	0,766
Variables de classe C			
EJT-Pas-Doul-Min (°)	0,216	0,056	0,043 !!
EJT-Pas-EDM-Min (°)	0,623	0,016 (g = 0,18)	0,060
FlxBassin-EDM (°)	0,872	<0,001 (g = 0,35)	0,117
FlxLomb-EDM (°)	0,907	0,436	0,643
FlxLatLomb-EDM-Min (°)	0,485	<0,001 (g = 0,31)	0,415
HanRI-Pas-EDM-Min (°)	0,223	0,576	0,545
HanRE-Pas-EDM-Min (°)	0,105	0,001 (g = 0,27)	0,721

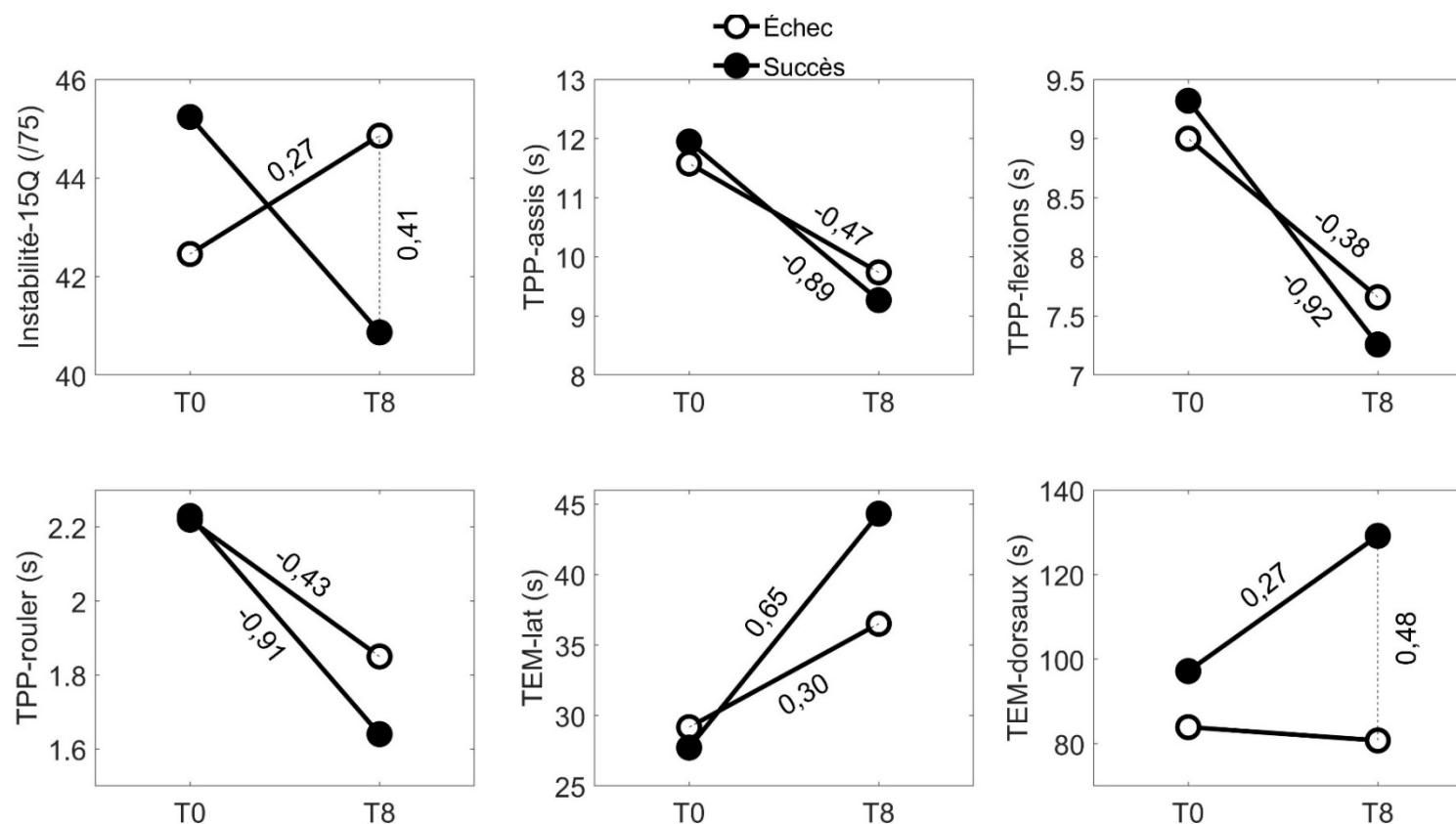
* Les variables qui ont été retenues dans la RPC ainsi que les valeurs $P \leq 0,05$ sont rédigées en caractères gras.

† Valeur positive lorsque GE > GS et négative lorsque GE < GS

‡ Les valeurs négatives indiquent une diminution dans le temps (entre T0 et T8).

!! Interaction illustrée dans la **figure 7**.

Figure 7. Illustration des interactions SOUS-GROUPE × TEMPS obtenues pour les variables PHT de type continu.



Les différences significatives obtenues avec les tests a posteriori sont identifiées avec les valeurs g correspondantes (tailles d'effet). TEM : test d'endurance musculaire ; TPP : test de performance physique.

Tableau 16. Effet du temps/traitement sur les mesures PHT (variables dichotomiques)

Variables*	Proportion (%) de tests positifs							
	Groupe échec (GE)				Groupe succès (GS)			
	T0	T8	T8-T0†	Valeur <i>P</i> (McNemar)	T0	T8	T8-T0†	Valeur <i>P</i> (McNemar)
Instabilité-Man	57	57	0	0,796	37	46	10	0,353
Instabilité-Ext	48	36	-11	0,166	44	21	-23	0,016
Mvt-Abe	18	18	0	0,739	40	13	-27	0,005
EJT-Act-Doul-Max	23	14	-9	0,405	48	12	-37	0,000
GenFlx-Pas-Max	34	23	-11	0,157	31	13	-17	0,013
GenFlx-Act-Max	34	23	-11	0,132	46	8	-38	0,000
HanRI-Pas-Max	57	34	-23	0,004	54	23	-31	0,003
HanRI-Act-Max	34	16	-18	0,021	37	12	-25	0,007
HanRE-Pas-Max	30	25	-5	0,593	40	8	-33	0,000
HanRE-Act-Max	34	30	-5	0,593	48	12	-37	0,000
HanExt-Pas-Max	41	45	5	0,683	35	8	-27	0,001
HanExt-Act-Max	82	50	-32	0,000	75	23	-52	0,000
ÉpaFlx-Act-Max	20	2	-18	0,011	35	8	-27	0,005
GenExt-Pas-Max	34	25	-9	0,248	52	23	-29	0,001
GenExt-Act-Max	34	20	-14	0,083	50	17	-33	0,002
HanRot-Pas-Max	20	18	-2	0,782	50	10	-40	0,000
HanRot-Act-Max	20	20	0	1,000	40	12	-29	0,009

* Les variables qui ont été retenues dans la RPC ainsi que les valeurs $P \leq 0,05$ sont en caractères gras.

† Changement de proportion (%) des patients ayant un test positif entre les temps T0 et T8; une valeur négative indique donc une amélioration.

4.4.2 Évaluation psychologique (mesures PSY)

4.4.2.1 Comparaisons entre sous-groupes et temps de mesure

Les ANOVA SOUS-GROUPE (succès par rapport à échec) $\times$ TEMPS (T0, T4, T8, T34) ont révélé six interactions (**tableau 17; figure 8**). Elles concernaient les peurs et croyances face à l'activité physique et la dramatisation de la douleur, la perception de la maladie, le sentiment d'efficacité personnel pour l'exercice (SEPE) et finalement l'intensité de la douleur et l'échelle STarT Back. À l'exception de SEPE, ces interactions se comportent toutes de la même façon, avec les GS et GE équivalents à T0. Une amélioration plus marquée des différents indicateurs est observée dans le temps chez le GS comparativement au GE, de sorte que les différences entre groupes, décelées dès T4, augmentent jusqu'à T8 pour finalement se stabiliser à T34 (**figure 8**). SEPE diminue aussi pour les deux groupes, mais de façon plus prononcée chez le GE ($g = -1,26$ à T34) comparativement au GS ($g = -0,61$ à T34), l'effet groupe étant également maximal à T8 ($g = 0,81$) et T34 ($g = 0,80$). D'ailleurs, l'effet temps était déjà détectable à T4 chez les GE ($g = -0,48$), contrairement au GS.

D'autres effets SOUS-GROUPE ont aussi été observés, mais sans interaction avec le facteur TEMPS, soient pour presque toutes les variables en lien avec la détresse psychologique (IDPcog, IDPanx, IDPdep, IDPcol, IDPsoma, IDPtot). Considérant l'absence d'interaction significative, ces indicateurs étaient plus favorables chez le GS dès le départ et pour tous les temps de mesure, mais les tailles d'effet étaient faibles (g entre 0,20 et 0,50).

L'effet du temps/traitement, sans interaction avec le facteur SOUS-GROUPE, a aussi été statistiquement significatif pour une majorité des mesures PSY continues investiguées (**tableau 17**), indiquant une amélioration dans le temps. Il s'agit de toutes les variables en lien avec la détresse psychologique (IDPtot et scores des cinq sous-échelles), des deux sous-échelles de soutien social/familial à l'exercice (SSESfamRP et SSESfamPa) et des attentes face au traitement (AttentesTrait), bien que les analyses a posteriori n'aient pas soutenu cet effet pour cette dernière. Les tailles d'effet relevées à T8 (par rapport à T0) étaient élevées (g entre -0,93 et -1,16) pour Peurs/croyances-AP, Dramatisation/douleur, PercMaladie.

La proportion des patients du GS ayant de la douleur lors de l'activité physique (DoulActPhys) a diminué avec le temps (77, 56, 25 et 23 % pour les temps T0, T4, T8 et T34 respectivement), une tendance qui semble plus favorable à T8 et T34, comparativement au GE (70, 59, 48 et 41 %), mais aucun test non paramétrique ne permet de faire cette comparaison à deux facteurs (SOUS-GROUPE $\times$ TEMPS). Bien que chacune des comparaisons (test de Cochran) ait mené à un effet principal TEMPS significatif ($P < 0,001$), les tests a posteriori (McNemar), combiné avec un ajustement de

l'alpha tenant compte du nombre de tests ($n = 6$, donc $P = 0,05/6 = 0,008$), n'ont été significatifs que pour GS ($T0 = T4$; $T0 < T8$ et $T34$; $T0 < T8$ et $T34$; $T8 = T34$).

Trois autres variables ont été recueillies aux temps T1, T4 et T8, soit la crédibilité (CEQ-crédibilité) et les attentes (CEQ-attentes) face au traitement ainsi que l'alliance thérapeutique entre le patient et son thérapeute (AllianceTrav). Aucune interaction SOUS-GROUPE $\times$ TEMPS n'a été observée (**tableau 18**). Ces variables sont interreliées avec des corrélations de Pearson variant entre 0,47 et 0,68 ($P < 0,001$), et démontrent toutes des valeurs significativement plus élevées chez le GS comparativement au GE, avec des tailles d'effet moyennes (0,61 à 0,71). La crédibilité et les attentes face au traitement ont augmenté ($P < 0,05$) entre T1 et T8, mais les tailles d'effets sont petites (0,20 et 0,25 respectivement).

Tableau 17. Effet de sous-groupe (SG) et du temps/traitement (T) sur les mesures PSY (variables continues) recueillies aux temps T0, T4, T8 et T34

Variables	Valeur <i>P</i> des ANOVA			Tests Post-hoc† (effet TEMPS)	Taille d'effet‡ (TEMPS) T8 - T0
	GROUPE (SG)	TEMPS (T)	SG × T		
	(taille d'effet*)	T0 vs T4 vs T8 vs T34			
Mesures PSY - Classe B					
Peurs/croyances- AP (/24)	0,002 (g = 0,41)	<0,001	0,021!!	T4, T8, T34	-0,93
Dramatisation/doul eur (/52)	<0,001 (g =0,51)	<0,001	0,001!!	T4, T8, T34	-0,95
IDPcog (/100)	0,041 (g = 0,20)	0,032	0,186	/	/
IDPanx (/100)	0,085	<0,001	0,435	T4, T8, T34	-0,37
IDPdep (/100)	0,005 (g = 0,30)	0,011	0,377	T8	-0,18
IDPcol (/100)	<0,001 (g =0,50)	0,010	0,131	T8	-0,21
IDPsoma (/100)	0,037 (g = 0,25)	<0,001	0,248	T4, T8, T34	-0,55
IDPtot (/100)	0,022 (g = 0,33)	<0,001	0,073	T4, T8, T34	-0,41
PercMaladie (/80)	<0,001 (g =0,91)	<0,001	<0,001!!	T4, T8, T34	-1,16
NAP-sport (/5)	0,560	0,274	0,050	/	/
NAP-loisirs (/5)	0,855	0,394	0,138	/	/
SEPE (/100)	0,028 (g = -0,43)	<0,001	0,048!!	T4, T34	/
SSFEfamRP (/100)	0,109	0,035	0,596	T34 (g = -0,36)	/
SSFEfamPa (/100)	0,879	<0,001	0,312	T34 (g = -0,45)	/
SSFEamiPa (/100)	0,787	0,381	0,249	/	/
AttentesTrait (1-5)	0,928	0,033	0,634	/	/
StadeChange (1-5)	0,944	0,075	0,610	/	/
Mesures PSY - Classe C					
Int-Douleur (/10)	<0,001 (g =0,73)	<0,001	<0,001!!	T4, T8, T34	-1,70
STarT Back (/9)	<0,001 (g =0,63)	<0,001	<0,001!!	T4, T8, T34	-1,26

* Valeur positive lorsque GE > GS et négative lorsque GE < GS

† Seulement les différences significatives par rapport à T0 sont identifiées.

‡ Les valeurs négatives indiquent une diminution dans le temps.

!! Interaction illustrée dans la **figure 8**.

Les valeurs $P \leq 0,05$ sont également en caractères gras.

Tableau 18. Effet du sous-groupe (SG) et du temps/traitement (T) sur les mesures PSY (variables continues) recueillies aux temps T1, T4 et T8

Variables	ANOVA (Valeur <i>P</i>)			Tests Post-hoc† (effet TEMPS)	Taille d'effet ‡ (TEMPS) T8 – T1
	GROUPE (SG) (taille d'effet*)	TEMPS (T) T1 vs T4 vs T8	SG × T		
CEQ-crédibilité (11-100)	<0,001 (g = -0,71)	0,002	0.418	T8	0,20
CEQ-attentes (11-100)	<0,001 (g = -0,61)	<0,001	0.080	T8	0,25
AllianceTrav (12-84)	0,002 (g = -0,64)	0,414	0.497	/	/

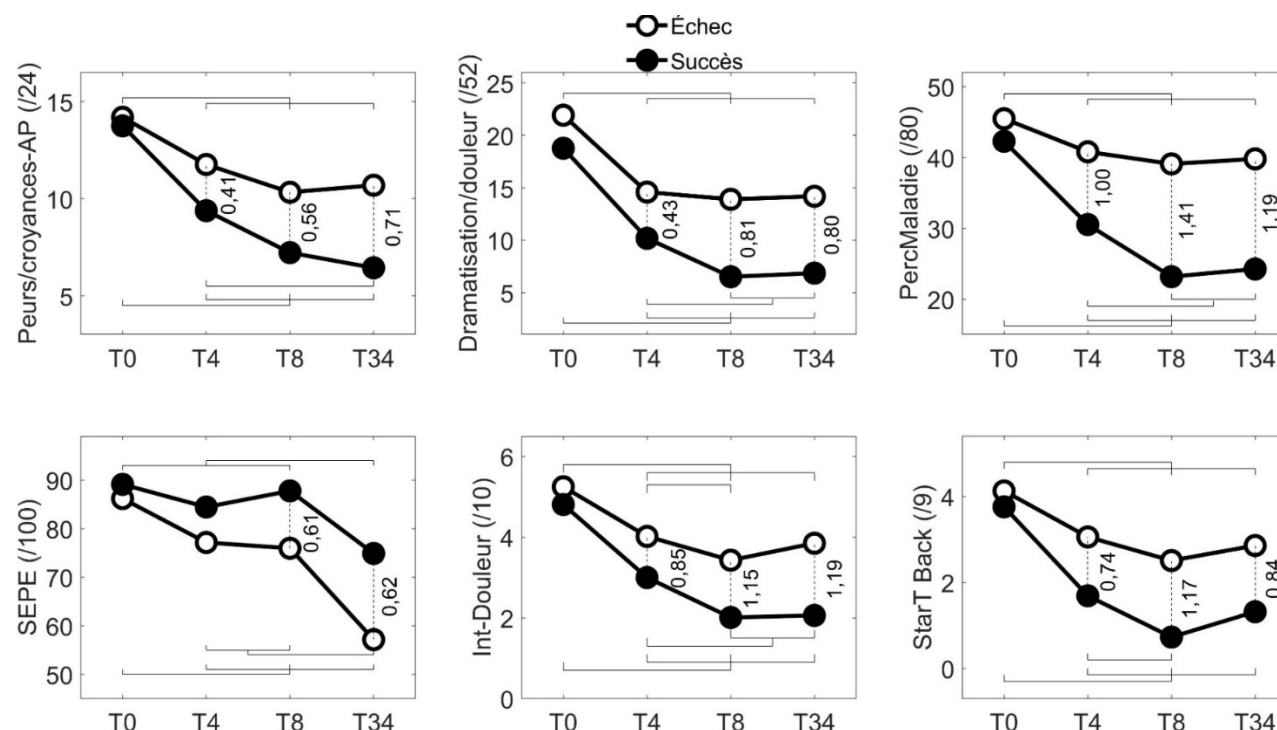
* Valeur positive lorsque GE > GS et négative lorsque GE < GS

† Seulement les différences significatives par rapport à T1 sont identifiées.

‡ Les valeurs négatives indiquent une diminution dans le temps.

Les valeurs $P \leq 0,05$ sont en caractères gras.

Figure 8. Illustration des interactions GROUPE × TEMPS obtenues pour les variables PSY de type continu.



Les différences significatives obtenues avec les tests *a posteriori* sont identifiées avec les valeurs *g* correspondantes (tailles d'effet) pour les différences entre groupes et par des lignes horizontales pour les différences entre temps de mesure.
 Peurs/croyances-AP : peurs et croyances face à l'activité physique; Dramatisation/douleur : dramatisation de la douleur ;
 PercMaladie : perception de la maladie ; SEPE : sentiment d'efficacité personnelle pour l'exercice ; Int-Douleur : intensité de la douleur ; STarT Back : Outil pronostic de dépistage.

4.4.2.2 Adhésion au programme d'exercice à domicile

L'analyse multivariée a permis de relever le rôle de l'interaction entre plusieurs variables (ChangGlob, OSW, SEPE, SSFEfamRP) et le temps pour expliquer l'adhésion à T8 et T34 (**tableau 19**). Ce modèle statistique multivarié est ajusté en fonction de quatre variables confondantes potentielles non modifiables, soient le sexe, l'âge, la durée de la douleur et le physiothérapeute responsable. La variance expliquée de ce modèle était de 76 %. ChangGlob était de loin le meilleur prédicteur selon les valeurs *P*. Étant donné qu'il s'agit d'une mesure agrégée intégrant de nombreux concepts, un second modèle de régression a été fait pour établir les variables qui y sont le plus associées. Toujours en ajustant pour les mêmes variables confondantes potentielles et toujours en interaction avec le temps, PercMaladie, SEPE (sentiment d'efficacité personnelle pour l'exercice), NAP-loisirs et OSW étaient les variables les plus importantes, avec une contribution marginale additionnelle de SSESfamRP (**tableau 20**). La variance expliquée de ChangGlob par ce modèle était de 50 %.

Tableau 19. Modèle multivarié de l'adhésion au programme d'exercice à T8 et T34

	Temps	Estimé	<i>P</i>	Intervalle de confiance à 95 %
ChangGlob × Temps	T34	0,1471	0,0007	[0,06627, 0,2279]
OSW × Temps	T34	0,009407	0,0325	[0,000821, 0,01799]
OSW × Temps	T8-T34	-0,01592	0,0182	[-0,02898, -0,00285]
SEPE × Temps	T8	0,004593	0,0493	[0,000016, 0,009170]
SEPE × Temps	T34	0,004315	0,0301	[0,000437, 0,008194]
SSESfamRP	T8-T34	-0,06102	0,0362	[-0,1179, -0,00413]

Tableau 20. Modèle multivarié du changement global (ChangGlob) perçu à T8 et T34

	Temps	Estimé	<i>P</i>	Intervalle de confiance à 95 %
PercMaladie × Temps	T8	-0,05495	<0,0001	[-0,07303, -0,03688]
PercMaladie × Temps	T34	-0,02717	0,0117	[-0,04752, -0,00683]
PercMaladie × Temps	T8-T34	-0,02778	0,0350	[-0,05338, -0,00218]
SEPE × Temps	T34	0,02163	0,0003	[0,01154, 0,03172]
SEPE × Temps	T8-T34	-0,02266	0,0072	[-0,03835, -0,00696]
NAP-loisirs × Temps	T34	-0,4874	0,0027	[-0,7825, -0,1924]
NAP-loisirs × Temps	T8-T34	0,4485	0,0199	[0,07982, 0,8171]
OSW × Temps	T34	-0,03100	0,0065	[-0,05217, -0,00983]
OSW × Temps	T8-T34	0,02998	0,0361	[0,002169, 0,05778]
SSESfamRP × Temps	T8-T34	-0,1559	0,0367	[-0,3011, -0,01073]

4.4.3 Évaluation neuromusculaire (mesures NRM)

Parmi les 110 patients qui ont complété le programme d'exercice et qui ont été évalués à T8, 77 patients ont fait l'objet d'une évaluation en laboratoire. Au temps T8, 31 se sont retrouvés dans le sous-groupe « échec », 6 dans le sous-groupe « amélioration clinique » et 40 dans le sous-groupe « succès », toujours selon les critères fixés précédemment avec les résultats du questionnaire Oswestry. Les sous-groupes échec et succès ont été comparés aux 30 sujets sains (15 hommes, 15 femmes) de la phase précédente. Rappelons qu'ils ont fait ces évaluations avec le même écart de temps (8 semaines), mais sans traitement. Les statistiques descriptives associées aux trois sous-groupes sont détaillées dans le **tableau 21**. Il y a apparence d'inégalité en nombre entre les hommes et les femmes entre le sous-groupe succès et les autres sous-groupes, mais le test Chi-carré est non significatif ($P = 0,314$) et confirme donc une représentation équilibrée des sexes à travers les sous-groupes. Les patients en échec de traitement avaient un IMC plus élevé que les sujets sains (**tableau 21**). Les femmes avaient un gabarit plus petit que les hommes ainsi qu'un pourcentage de graisse plus élevé.

Avant de procéder aux analyses sur les mesures NRM, des analyses (ANOVA avec mesures répétées) ont été réalisées afin de connaître l'effet du programme d'exercice, chez les 71 patients, en ce qui a trait aux mesures de résultats primaires (OSW, Int-Douleur, ChangeGlob) (**tableau 22**). Comme pour l'échantillon complet de patients, les effets sont visibles sur toutes les mesures de résultats primaires, notamment les effets plus importants observés pour le groupe avec succès, comme le démontrent les interactions GROUPE × TEMPS significatives pour les variables OSW et Int-Douleur ainsi que l'effet GROUPE pour la variable ChangeGlob.

Les analyses ANOVA ont été très peu révélatrices en ce qui concerne toutes les mesures neuromusculaires (proprioception lombaire, équilibre postural en position assise sur chaise instable, rigidité lombaire, ajustements posturaux anticipatoires, coordination du tronc lors de la flexion, mesures ultrasonographiques). Il y a eu quelques interactions SOUS-GROUPE × TEMPS et des effets TEMPS statistiquement significatifs, mais qui se sont soldés par des effets négligeables sur le plan clinique (petites tailles d'effet). De nombreux effets GROUPES ont été observés, notamment pour les muscles ultrasonographiques, mais ces effets n'étaient pas au cœur du questionnement de la présente étude, à savoir si le PESL avait des effets sur les mesures neuromusculaires. Pour ces raisons, les détails des résultats ne sont pas présentés dans le présent rapport.

Tableau 21. Description des caractéristiques démographiques et anthropométriques des sujets sains et lombalgiques qui ont réalisé les tests neuromusculaires

Variables	Sujets sains (n = 30)				Sujets lombalgiques avec échec de traitement (n = 31)				Sujets lombalgiques avec succès de traitement (n = 40)				Valeur P (ANOVA)		
	Hommes (n = 15)		Femmes (n = 15)		Hommes (n = 17)		Femmes (n = 14)		Hommes (n = 15)		Femmes (n = 25)		GROUPE (G)	SEXE (S)	G×S
	M	(ET)	M	(ET)	M	(ET)	M	(ET)	M	(ET)	M	(ET)			
Âge (années)	39	(14)	40	(14)	46	(12)	40	(13)	42	(13)	43	(13)	0,289	0,452	0,456
Taille (m)	1,78	(0,08)	1,64	(0,06)	1,73	(0,06)	1,64	(0,06)	1,77	(0,07)	1,63	(0,06)	0,402	< 0,001	0,669
Masse (kg)	77,1	(10,7)	62,9	(10,6)	83,0	(11,5)	67,5	(6,3)	76,7	(12,9)	67,3	(10,7)	0,159	< 0,001	0,584
IMC (kg/m ²)	24,3	(3,2)	23,3	(3,6)	27,3	(3,2)	25,3	(2,5)	24,5	(2,9)	25,2	(3,3)	0,012	0,340	0,251
% graisse	21	(7)	33	(6)	25	(6)	33	(6)	21	(5)	33	(5)	0,461	< 0,001	0,254

M : moyenne; ET : écart-type; IMC : indice de masse corporelle; % graisse : estimé à partir des plis cutanés (biceps, triceps, abdominal, sous-scapulaire) (Durnin et Womersley, 1974); Les valeurs $P \leq 0,05$ sont en caractères gras alors que les valeurs situées entre 0,05 et 0,10 sont soulignées.

Tableau 22. Description des mesures de résultats (OUT) des sujets lombalgiques qui ont réalisé les tests neuromusculaires

Variables	Temps	Sujets lombalgiques en échec				Sujets lombalgiques en succès				Valeur P (ANOVA)*					
		♂ (n = 17)		♀ (n = 14)		♂ (n = 15)		♀ (n = 25)		Groupe	Sexe	Temps	G×S	G×T	S×T
		M	(ET)	M	(ET)	M	(ET)	M	(ET)						
Oswestry (%)	T0	22	(8)	24	(9)	28	(8)	28	(10)	0,002	0,287	< 0,001	0,474	< 0,001	0,594
	T8	19	(9)	23	(9)	6	(4)	7	(5)						
Int-Douleur (/10)	T0	5,1	(1,5)	5,1	(1,1)	5,0	(1,5)	4,7	(1,0)	< 0,001	0,551	< 0,001	0,747	< 0,001	0,787
	T8	3,3	(1,1)	2,9	(1,1)	1,9	(1,2)	1,9	(0,8)						
ChangGlob (-3 à 3)	T0	/	/	/	/	/	/	/	/	< 0,001	<u>0,058</u>	/	0,849	/	/
	T8	1,5	(0,7)	1,9	(1,0)	2,2	(0,8)	2,6	(0,6)						

♂ : homme ; ♀ : femme ; M : Moyenne; ET : écart-type. Les valeurs $P \leq 0,05$ sont en caractères gras alors que les valeurs situées entre 0,05 et 0,10 sont soulignées. *Les interactions triples (G×S×T) sont toutes non significatives. * Pour la variable ChangGlob, une ANOVA à deux voies a été réalisée pour évaluer les effets GROUPE et SEXE étant donné que cette mesure ne pouvait être prise qu'à T8.

5. DISCUSSION

En résumé, la présente étude a mené : (1) à la dérivation de deux RPC de succès ayant toutes deux une bonne valeur prédictive, soient pour prédire le succès à T8 (fin du traitement) et T34 (suivi de six mois) ; (2) à la conclusion que plusieurs mécanismes physiques et psychologiques ont le potentiel d'expliquer le succès au PESL. Ceci n'était pas le cas pour les mécanismes de nature neuromusculaire. Finalement, l'étude plus spécifique de l'adhésion au programme d'exercices à domicile suggère que certains des mécanismes de nature psychologique en question joueraient un rôle dans cette adhésion, ce qui à son tour aurait un impact sur le succès du PESL.

Le programme d'exercice a produit des effets cliniques importants sur l'ensemble du groupe ($n = 110$), avec des d de Cohen de $-1,24$ et $-1,70$, correspondant respectivement à la diminution de OSW et Int-Douleur à T8. Les scores correspondants étaient de $-1,24$ et $-1,32$ au suivi de six mois (T34). Les d de Cohen dépendent beaucoup des critères d'inclusion utilisés, notamment lorsqu'un seuil minimal d'OSW est requis. Dans les études où l'information nécessaire était disponible et ayant recruté une clientèle de patients en phase chronique, ces valeurs d variaient entre $-0,58$ et $-1,11$ (incapacités) et entre $-0,61$ et $-1,02$ (Int-Douleur) lorsqu'aucun critère de sévérité (selon OSW) n'était utilisé (Ferreira *et al.*, 2007; Goldby *et al.*, 2006) alors qu'elles sont plus élevées (OSW : $-1,99$; Int-Douleur : $-1,43$) lorsqu'un critère d'incapacité était utilisé (Rabin *et al.*, 2014). Les effets observés dans la présente étude étaient donc relativement importants par rapport à ce qui est rapporté dans la littérature, ce qui s'expliquerait par le recrutement d'une clientèle plus sévèrement atteinte que dans des études antérieures, bien que le score OSW moyen de 27 % demeure relativement faible.

5.1 Objectif 1 – Dérivation des RPC de succès

Dans la section qui suit, nous avons choisi de mettre en exergue les différentes comparaisons possibles avec les RPC préliminaires déjà dérivées par certains auteurs, notamment la RPC référencée par Hicks *et al.* (2005). Par contre, il faut souligner que l'étude de Hicks se différencie de la présente étude par la population étudiée, car les patients en phase aiguë étaient aussi recrutés, ce qui a mené à un échantillon ayant une durée moyenne des symptômes de 41 jours, ce qui est nettement inférieur à celle de notre échantillon composé à 97 % (107/110) de patients avec douleurs chroniques (3 mois ou plus). Aussi, bien que cette question soit toujours controversée, rappelons que le PESL de Hicks *et al.* (2005) était basé sur l'école de pensée de McGill (2002), impliquant une cocontraction globale des muscles du tronc, alors que le présent PESL est basé sur une approche hybride combinant l'approche australienne (Richardson *et al.*, 2004), qui se concentre sur le contrôle isolé des muscles profonds du tronc, suivi par l'approche de McGill pour rehausser l'endurance musculaire.

Prenant en compte la nature des variables constituant nos modèles, la discussion de cette section (objectif 1) est d'abord divisée en trois parties, suivant la force de la relation théorique entre les variables considérées dans les modèles et l'instabilité lombaire : (1) modèles avec variables de classe A seulement ; (2) modèles avec variables de classe A et B et (3) modèles avec variables de classe A, B et C. Finalement, une quatrième sous-section discutera d'un point abordé pour la première fois dans cette littérature, portant sur le choix de la RPC la plus utile, à savoir si elle devrait prédire le succès à la fin du programme clinique ou le succès après un suivi de six mois.

5.1.1 Modèles ne comportant que des variables de classe A

Les modèles 3 et 5 à T8 (**tableau 7**) et 2 et 3 à T34 (**tableau 11**) ne comportaient que des variables de classe A, dont deux variables qui avaient préalablement été identifiées dans la RPC préliminaire ne portant que sur un effectif réduit, soient TPP-portée et HanRot-Pas-Max. Ceci prouve la robustesse de leur contribution dans le pouvoir explicatif du modèle.

Sur la base de l'indicateur statistique le plus important de toute RPC de succès (LR+), la valeur prédictive globale de ces quatre modèles, comportant entre 3 et 5 variables, variait entre 7,5 et 8,3 (**tableaux 8 et 12**), ce qui devance largement le modèle préliminaire (LR+ = 4,0) de Hicks *et al.* (2005) de même que notre propre modèle préliminaire (Larivière *et al.*, 2016) offrant un LR+ de 1,9. Rappelons que la RPC de succès de Hicks *et al.* (2005) comprenait quatre variables : (1) test d'instabilité lombaire positif (notre variable Instabilité-Man), (2) présence d'au moins un des cinq signes de mouvements aberrants lors d'un mouvement de flexion/extension maximale du tronc (notre variable Mvt-Abe), (3) test élévation passive jambe tendue > 91° (notre variable EJT-Pas-Max-EDM) et (4) âge < 40 ans. Notons la présence d'une variable de classe C (EJT-Pas-Max-EDM) dans cette RPC.

Rappelons que les RPC finales sélectionnées (modèle 3 à T8 et modèle 2 à T34) ne comportaient que des variables de classe A, ce qui permet de comparer leurs performances globales dans la présente section. Leur étude plus approfondie en fonction du nombre de tests positifs (**tableaux 10 et 14**) révèle une performance équivalente à T8 (LR+ = 17,9; considérant 2 prédicteurs ou plus) et à T34 (LR+ = 17,0). Ces performances, de même que les probabilités de succès, sont comparées aux différentes RPC de succès développées pour les traitements offerts en physiothérapie en lien avec la colonne vertébrale (**tableau 23**). Les RPC à T8 et T34 se positionnent bien comparativement à ces autres RPC, notamment vis-à-vis la RPC de succès à T8 dérivée par Hicks *et al.* (2005) pour les exercices de stabilisation lombaire (LR+ = 4,0; probabilité de succès de 67%).

Tableau 23. Comparaison entre les performances des différentes RPC de succès développées pour différents traitements offerts en physiothérapie.

Étude dérivation RPC de succès	Traitement	LR+	Probabilité de succès sans/avec RPC (différence)
Présente étude – à T8	Exercices de stabilisation lombaire	17,9	49 / 96 % ($\Delta = +47 \%$)
Présente étude – à T34	Exercices de stabilisation lombaire	17,0	53 / 85 % ($\Delta = +32 \%$)
Hicks <i>et al.</i> (2005)	Exercices de stabilisation lombaire	4,0	33 / 67 % ($\Delta = +34 \%$)
Stolze <i>et al.</i> (2012)	Exercices de Pilates	10,6	54 / 93 % ($\Delta = +39 \%$)
Cai <i>et al.</i> (2009)	Traction de la colonne vertébrale	9,4	19 / 69 % ($\Delta = +50 \%$)
Cleland <i>et al.</i> (2007)	Manipulation thoracique	5,5	54 / 86 % ($\Delta = +32 \%$)
Flynn <i>et al.</i> (2002)	Manipulation spinale	24,4	45 / 95 % ($\Delta = +50 \%$)

Afin de nous aider à juger de l'interprétation clinique des RPC, voici l'échelle qualitative d'appréciation de l'indicateur LR+ (Jaeschke *et al.*, 1994): **(1) LR+ > 10** génèrent des changements importants et souvent concluants entre la probabilité en prétraitement à celle en post-traitement; **(2) LR+ de 5 à 10** génèrent des changements modérés entre la probabilité en prétraitement et post-traitement; **(3) LR+ de 2 à 5** génèrent des changements de probabilité faibles (mais parfois importants); **(4) LR+ de 1 à 2** modifient la probabilité à un faible degré (et rarement important). Conséquemment, les RPC à T8 et T34 figurent toutes deux dans la catégorie supérieure ($LR+ > 10$), suggérant des changements importants et souvent concluants dans la probabilité de succès. En effet, selon les estimations (**tableau 23**), cela semble particulièrement le cas pour la RPC à T8 (de 49 à 96 %, soit un changement de 47 %), mais un peu moins pour la RPC à T34 (de 53 à 85 %, représentant un changement de 32 %).

Rappelons que les variables de classe A ont toutes un lien théorique, mais potentiellement inégal, avec l'instabilité lombaire. Sur la base de leur poids relatif, tel qu'inféré avec la statistique adéquation (**tableau 7** à T8; **tableau 11** à T34), il est possible de mieux jauger leur poids relatif, pour les lecteurs intéressés. Par contre, ce poids relatif n'a pas d'influence sur la décision que prendra le thérapeute, car au final, c'est le nombre de prédicteurs positifs (ou rencontrant le seuil indiqué) qui est important à connaître pour la prise de décision, ce qui dans les faits redonne un poids équivalent entre ces prédicteurs. Ces prédicteurs seront donc discutés sans égard à leur poids relatif initial dans les prochains paragraphes. Par contre, il sera fait mention si elles ont été sélectionnées dans les RPC finales afin d'aider le lecteur à dégager l'information la plus pratique (RPC finales pour les thérapeutes) de l'information plus théorique (prédicteur potentiel à suivre par la communauté scientifique).

Le test de performance physique **TPP-portée** a été sélectionné dans la RPC finale à T8, ceci avec un seuil $\geq 0,75/1$. Ce test semble avoir des assises théoriques en matière d'instabilité lombaire. Certaines caractéristiques de ce test permettent de mettre à l'épreuve le rôle des muscles stabilisateurs de la colonne lombaire. Premièrement, la stabilité lombaire est moindre lorsque la région lombaire est en position neutre (comme en

position debout), car les tissus passifs ne peuvent contribuer à stabiliser les articulations lombaires (Cholewicki et McGill, 1996). Deuxièmement, le fait d'élever la charge à la hauteur des épaules augmente aussi l'instabilité lombaire, comme le démontre l'augmentation de l'activation des muscles du tronc lorsque la charge est plus haute (Ghezelbash *et al.*, 2021; Granata et Orishimo, 2001). Finalement, une réaction similaire est observée lorsque la charge soutenue est maintenue loin du corps, car non seulement l'activation des muscles dorsaux est augmentée pour contrer le moment net en flexion, mais il y a aussi augmentation de l'activation des muscles abdominaux qui ne peut être attribuable qu'au maintien de la stabilité lombaire (Ghezelbash *et al.*, 2021; Larivière, Shahvarpour, *et al.*, 2019). Notre hypothèse de départ était que les patients qui stabilisent moins bien leur colonne lombaire auraient de la difficulté à réaliser ce test (scores inférieurs à 0,75/1) et bénéficieraient du programme d'exercice de stabilisation lombaire. Or, ce sont les patients avec des scores supérieurs qui semblent pouvoir en bénéficier. L'explication possible serait que l'éloignement de la charge, induisant une cocontraction automatique des muscles du tronc, augmenterait la stabilité lombaire et réduirait la douleur des patients dont la douleur est provoquée par une instabilité articulaire. Ces patients bénéficieraient donc d'un programme d'exercice de stabilisation lombaire pour apprendre à mieux se servir de cette cocontraction lors de différentes tâches de plus en plus fonctionnelles. Les patients qui ont une douleur qui ne se contrôle pas avec la cocontraction des muscles du tronc performeraient donc moins bien lors de ce test et, par conséquent, ne bénéficieraient pas de ce programme d'exercice. Il s'agit de la même théorie sous-jacente au test d'instabilité en position décubitus ventral (*prone instability test*, soit notre variable Instabilité-Man). En effet, ce test est positif lorsque le patient voit sa douleur lombaire réduite (ou disparaître selon les auteurs et comme ce fut le cas dans la présente étude) lorsqu'il soulève les membres inférieurs du sol, car ce mouvement a aussi pour effet d'induire une cocontraction des muscles du tronc et une meilleure stabilité lombaire. Cela pourrait donc expliquer pourquoi TPP-portée ressort dans notre RPC, à la place d'Instabilité-Man qui ressortait dans celle de Hicks *et al.* (2005). Si TPP-portée mesure vraiment le même phénomène, ce test représenterait une alternative intéressante, bien que du matériel soit nécessaire pour le réaliser (sac avec poids et ruban à mesurer). D'ailleurs, la mesure continue (en cm) a une fidélité interjuges quasi parfaite (CCIC : 0,99) et une excellente fidélité interjours (2 semaines d'écart) (CCIC : 0,91) (Simmonds *et al.*, 1998), ce qui devrait aussi se traduire positivement par des scores dichotomisés. Cette bonne fidélité n'est peut-être pas étrangère au fait que ce test ne requière pas d'intervention manuelle de la part du thérapeute. Le test d'instabilité en position décubitus ventral (variable Instabilité-Man) fait l'objet de controverse (Hicks, 2011) au regard de sa fidélité interjuges, car elle s'avère très variable (kappa entre 0,27 et 0,87) d'une étude à l'autre (Stolz *et al.*, 2019), ce qui pourrait être attribuable au type de patients recrutés dans ces études (lombalgie récurrente vs en phase aiguë vs chronique) selon les auteurs de cette revue. Hicks (2011) évoque plutôt un suivi plus strict des procédures entourant la standardisation de ce test. Une étude plus récente de l'activation des différents muscles ayant le potentiel d'augmenter la rigidité lombaire lors de ce test (lors de l'élévation des

membres inférieurs) démontre des patrons d'activation (synergies) différents entre des participants avec (n = 10) et sans (n = 10) douleurs lombaires (Sung *et al.*, 2019). Il n'est donc pas exclu que des patrons d'activation différents puissent être associés à différents sous-groupes de patients, ce qui en retour aurait le potentiel de faire varier les symptômes lors de ce test et même lors du test TPP-portée. Des études de ce type doivent donc être menées chez un plus grand échantillon de participants pour vérifier cette possibilité.

La variable **HanExt-Act-Max** (RPC finale à T34) est issue d'un test de déficience du contrôle sensorimoteur (en position décubitus ventrale) demandant une extension active de la hanche suffisante pour que la cuisse ne touche plus la table d'examen. Les mouvements des membres, notamment lorsqu'un seul membre est impliqué, provoquent généralement des forces (et mouvements) asymétriques au niveau lombaire, générant une rotation axiale. Il est proposé qu'une inhabilité à limiter ces mouvements serait révélatrice d'une instabilité lombaire, provoquant ainsi des symptômes (Liebenson *et al.*, 2009). Or, dans le présent cas, le test doit être négatif pour prédire le succès, ce qui signifie que le mouvement ne doit rien changer ou bien doit réduire les symptômes par rapport à la position de référence. Il est possible que les patients qui démontrent déjà une meilleure fonction des muscles du tronc, sur au moins certaines mesures, puissent bénéficier plus rapidement du PESL (8 semaines seraient suffisantes) et/ou soient plus susceptibles de conserver ces avantages fonctionnels plus longtemps après l'arrêt du PESL (c'est-à-dire lors du suivi à 6 mois). Ce résultat semble être cohérent avec les autres facteurs prédictifs qui donnent des scores plus élevés, tel que pour TPP-Portée et les mesures d'endurance des muscles du tronc (TEM-Dorsaux et TEM-Abdominaux ; discuté ci-dessous).

La variable **HanRot-Pas-Max** (RPC finale à T8) est issue d'un test de déficience du contrôle sensorimoteur où il y a abduction et rotation passives de la hanche. L'association entre cette variable et l'instabilité lombaire n'est pas évidente parce que le test est passif. Ce test de provocation de la douleur peut ne pas être spécifique aux patients qui ont besoin d'un programme d'exercice de stabilité de la colonne lombaire pour éliminer ces symptômes. Il n'est pas non plus exclu que ce test ait provoqué une irritation de l'articulation sacro-iliaque qui se retrouve sous tension lors de ce test, car la présence de douleur d'origine sacro-iliaque ne représentait pas un critère d'exclusion dans la présente étude. C'est à l'étape de validation de la RPC qu'il sera possible d'établir si ce test est spécifique ou non à un trouble d'instabilité lombaire.

La présence de **mouvements aberrants (Mvt-Abe)** (RPC finale à T8) avait aussi été retenue dans la RPC de succès de Hicks *et al.* (2005). Il est aussi intéressant de noter que la tentative de validation de la RPC de Hicks *et al.* (2005), par Rabin *et al.* (2014), a mener à la proposition d'une nouvelle RPC de succès plus prédictive que la version initiale, ne retenant que les mouvements aberrants et le test d'instabilité en position décubitus ventral (*prone instability test*). Les mouvements aberrants sont associés à une diminution du contrôle de la mobilité des segments lombaires dans la portion centrale

(c.-à-d. ni au début ni à la fin) de l'étendue de mouvement, comme démontré par l'étude fluoroscopique (technique d'imagerie dynamique) des mouvements de flexion et d'extension dans le plan sagittal (Teyhen *et al.*, 2007). Rappelons que ce test est basé sur un ensemble d'observations ($n = 5$) et que la seule façon d'obtenir une fidélité interévaluateurs acceptable ($k = 0,60$ à $0,65$) est de considérer le test positif lorsqu'une seule de ces observations s'avère anormale. Des études plus approfondies sur les mouvements aberrants supportent la validité de construit de ce test, mais fait aussi remarquer que cette façon d'opérationnaliser ce test ne permet pas de distinguer des sous-groupes de patients ayant des pathologies différentes (Biely *et al.*, 2014).

La **présence d'une laxité ligamentaire généralisée** (variable **Beighton** $< 5/9$) a été retenue dans la RPC finale à T8, bien que l'échelle de Beighton ne soit pas spécifique à la région lombaire. Par contre, la direction de la relation est contre-intuitive. En effet, ce sont les patients avec un score élevé (et non faible) sur l'échelle Beighton qui devraient théoriquement être associés à une instabilité lombaire et qui devraient bénéficier du programme d'exercices de stabilisation lombaire. Une tentative d'explication, hautement spéculative, serait que les personnes ayant un haut score de Beighton, présumèrent depuis la naissance, auraient appris à contrôler leurs mouvements dans cette condition défavorable et n'auraient donc pas besoin d'un PESL. Le syndrome d'hypermobilité articulaire, tel que diagnostiqué avec cette échelle, serait causé par une diminution de la teneur en collagène dans les tendons, ligaments, capsules articulaires et la peau, ce qui en retour réduit leur rigidité et peut donc avoir une relation avec l'instabilité articulaire et la douleur (van Meulenbroek *et al.*, 2020). Une relation avec le sexe est attendue (Remvig *et al.*, 2007a), tel qu'il a été observé dans notre échantillon avec un score supérieur (test de Wilcoxon; $P < 0,001$) chez les femmes ($2,2 \pm 2,6$; $n = 53$) que chez les hommes ($0,5 \pm 1,4$ ($0,5 \pm 1,4$; $n = 46$)). Le seuil de $\geq 4/9$ est le plus utilisé pour établir le diagnostic, bien que des seuils de $\geq 5/9$ ou $\geq 6/9$ aient aussi été proposés (Remvig *et al.*, 2007a), ce qui supporte le seuil qui a été dérivé dans la présente étude à partir d'un critère qui n'avait jamais été utilisé auparavant (succès à un PESL) (Remvig *et al.*, 2007b). Une revue systématique de la littérature a conclu que ces personnes présentent également une moins bonne proprioception (Smith, T. O. *et al.*, 2013), ce qui a le potentiel d'accentuer l'instabilité articulaire. Finalement, une récente revue de la littérature apporte des arguments pour formuler une hypothèse établissant un lien entre ce syndrome d'hypermobilité articulaire et l'incapacité, via les processus du modèle d'évitement de la peur (van Meulenbroek *et al.*, 2020). En effet, cette hypermobilité, qui serait à la source de douleur, entraînerait les peurs et l'évitement qui à leur tour réduiraient l'exposition aux mouvements et à l'entretien des mécanismes proprioceptifs et d'équilibre postural nécessaires à la stabilité articulaire. Par contre, dans la présente étude ($n = 99$ patients à T8), la laxité ligamentaire n'était pas associée aux variables psychologiques associées à la douleur (Peurs/croyances-AP, Dramatisation/douleur), ni au niveau d'activité physique (NAP-sport, NAP-loisirs), mais était faiblement associée à la présence de mouvements aberrants ($r = 0,21$; $P = 0,035$).

L'endurance des muscles dorsaux (variable TEM-dorsaux > 225 s) n'a pas été retenue dans la RPC finale à T8 (présente dans les autres modèles), mais **l'endurance des muscles abdominaux (variable TEM-abdominaux > 72,62 s)** a été retenue dans la RPC finale à T34. Il est intéressant de noter que le modèle 6 à T8 (pas présenté ici) excluait **TEM-dorsaux** pour la remplacer par **TEM-abdominaux** et l'IMC, mais ces deux variables n'avaient tout juste pas atteint le niveau de signification statistique ($P = 0,064$ et $0,065$, respectivement). De tels résultats n'ont pas été observés dans l'étude de Hicks *et al.* (2005), possiblement à cause de leur échantillon composé de patients avec une durée des symptômes bien moindre que dans la présente étude. En effet, l'endurance des muscles du tronc est corrélée négativement avec la durée des symptômes (Vanti *et al.*, 2016), ce qui s'expliquerait par une transformation progressive des fibres musculaires en fibres plus « fatigables » (Mannion *et al.*, 2000). Ces prédicteurs seraient liés à l'effet protecteur que jouerait l'endurance sur la stabilité lombaire, comme promue par l'école de pensée de McGill (2002). Par contre, la direction de la relation avec le succès du PESL est contre-intuitive, car les patients avec une moins bonne endurance devraient en bénéficier et non le contraire. Une tentative d'explication serait que les patients avec instabilité lombaire auraient une inhibition des muscles profonds (LuM, TrA), comme le soutiennent certaines études, ce qui en retour les amènerait à surutiliser les muscles superficiels (van Dieën *et al.*, 2017) et a ainsi développé l'endurance de ces muscles. Or, les muscles superficiels disposent d'un avantage mécanique (bras de levier plus longs) par rapport aux muscles profonds, ce qui pourrait mener à de meilleures performances lors de tels tests d'endurance. Cette hypothèse dispose de quelques éléments de support. Des patients démontrant une instabilité clinique, c.-à-d. répondant positivement à la RPC de succès de Hicks *et al.* (2005), ont démontré une plus grande endurance des muscles dorsaux lors d'une épreuve d'endurance basée sur un effort relatif (60 % de la force maximale) (Davarian *et al.*, 2012), ce qui est différent des épreuves de support du poids du tronc comme fait dans la présente étude, les deux types d'épreuves pouvant potentiellement introduire des biais, comme discuté par ces auteurs. Une seconde étude a démontré des corrélations positives entre un test d'endurance des muscles dorsaux (prone bridge test) et différents tests d'instabilité lombaire tels que réalisés dans la présente étude (Mvt-Abe, Instabilité-Man, Instabilité-Ext, EJT-Act-Doul) (Vanti *et al.*, 2016), bien que ces auteurs semblent avoir mal interprété le sens de ces corrélations, suggérant qu'une moindre endurance serait attendue chez les patients démontrant des signes d'instabilité lombaire. Une explication alternative serait que ces prédicteurs ne sont pas spécifiques à ce traitement et ne représentent que des facteurs pronostiques. En effet, une pauvre endurance des muscles dorsaux prédit un premier épisode de lombalgie (Biering-Sorensen, 1984) ainsi qu'une incapacité prolongée (Enthoven *et al.*, 2003). Dans le présent PESL, l'emphasis placée sur le développement de l'endurance des muscles du tronc serait justifiée selon ces résultats. Cependant, la mesure de TEM-dorsaux pour le modèle 5 mènerait à des performances quasi équivalentes au modèle 3 en termes de prédiction du succès (**tableau 7**), justifiant l'élimination de ce test à l'étape de l'examen initial servant à déterminer le traitement à venir. En effet, ceci réduit considérablement le

temps d'évaluation et sa pénibilité pour le patient. Par contre, dans le cas où la RPC de succès (modèle 3) mènerait à choisir le PESL comme modalité de traitement, il apparaîtrait justifié de mesurer l'endurance des différents muscles du tronc (TEM-dorsaux, TEM-abdo, TEM-lat) à titre d'indicateurs de l'évolution du patient au cours du PESL.

Le temps pris pour exécuter cinq flexions/extensions du tronc (variable TPP-flexions $\geq 9,91$ s) a été retenu dans la RPC finale à T34. Il s'agit d'un test de performance qui peut être lié à l'instabilité lombaire, car les muscles du tronc doivent plus s'impliquer lorsque la colonne vertébrale lombaire est soumise à des perturbations (initiées volontairement dans cette tâche), ceci dans le but d'augmenter la rigidité lombaire et ainsi, la stabilité lombaire. Ceci est vrai pour toutes les personnes ayant une lombalgie (Lamoth *et al.*, 2006), mais c'est possiblement plus critique chez les personnes dont la lombalgie est expliquée par une instabilité lombaire. Or, si l'on tient pour acquis que les participants qui ont eu un succès avec le PESL avaient une instabilité lombaire, il serait normal pour eux de diminuer la vitesse d'exécution de cette tâche, notamment dans la portion où le tronc est à la verticale. En effet, non seulement la stabilité lombaire est moindre lorsque la colonne lombaire est dans la zone neutre (ou verticale) comparativement à la position fléchie, tel qu'estimé avec des modèles biomécaniques internes de la colonne lombaire (Arjmand et Shirazi-Adl, 2006; Cholewicki *et al.*, 1997; Shahvarpour *et al.*, 2015), mais cette tâche impose aussi un changement de direction (de l'extension vers la flexion) dans cette posture plus précaire, et ceci à quatre reprises. Ces changements de direction imposent des défis supplémentaires en termes de stabilité lombaire. En dehors des estimations produites par les modèles biomécaniques, une approche conceptuelle très différente permettant l'étude de la stabilité lombaire (nommé stabilité dynamique locale), laquelle est basée sur l'étude du mouvement lombaire lors de mouvements répétés de flexion/extension du tronc. Or, cette méthode permet d'arriver aux mêmes conclusions, à savoir qu'il est préférable de réduire la vitesse de mouvement pour améliorer la stabilité lombaire (Asgari *et al.*, 2015; Granata et England, 2006).

Finalement, l'âge (< 43 ans) a été retenu dans la RPC finale à T34, ce qui est en accord presque parfait avec la RPC de succès de Hicks *et al.* (2005), le seuil étant légèrement plus bas (âge < 40 ans). Bien que l'âge ait été considéré comme une variable de classe A, le lien avec l'instabilité lombaire est à déterminer, si tel est vraiment le cas. Pour l'instant, l'âge, au même titre que le sexe ou l'IMC, a été considéré pour un ajustement possible des modèles à des variables confondantes potentielles.

5.1.2 Modèles comportant des variables de classe A et B

Lorsque des variables de classe B étaient admissibles, en plus des variables de classe A, deux variables ont été introduites autant à T8 qu'à T34, soit la perception de la maladie (PercMaladie) mesurée avec le bref questionnaire (8 items) portant sur les perceptions de la maladie (Broadbent *et al.*, 2006) et les peurs et croyances associées à l'activité physique (Peurs/croyances-AP) mesurées avec la sous-échelle (5 items) de ce

questionnaire (Waddell *et al.*, 1993). Tout comme les variables de classe C, les variables de classe B sont également non spécifiques au PESL, mais ont le potentiel d'être associées à son succès par l'entremise de l'adhésion au programme d'exercices fait à domicile. Les analyses multivariées pour prédire l'adhésion au PESL à domicile ont justement permis de faire, entre autres, le lien avec notre mesure de résultat principale d'incapacités perçues (OSW). Ce modèle multivarié et le devis de l'étude ne permettant pas d'établir la direction et la causalité de cette relation. Théoriquement, les deux interprétations sont possibles. En effet, une réduction de l'incapacité (OSW) pourrait contribuer à renforcer l'adhésion au programme, tout comme, le fait d'adhérer au programme peut avoir contribué à réduire l'incapacité, de par l'efficacité du traitement. Finalement, une relation causale bidirectionnelle est aussi possible.

La perception/compréhension de la maladie (PercMaladie), avec un score $\leq 38,5 / 80$ (modèles à T8) ou $< 44/80$ ou $45/80$ (modèles à T34), identifierait une compréhension juste ou du moins non polarisée de sa condition de santé. Elle prédirait aussi le succès du PESL. Donc, les patients ayant un score plus élevé que ces seuils auraient une vision plus polarisée de leur condition, correspondant à une perception/compréhension en partie erronée de sa condition de santé. Ceux-ci ne bénéficieraient pas autant du PESL. Ces incompréhensions pourraient expliquer une moins grande adhésion au PESL fait à domicile. Théoriquement, ce résultat est logique puisque le modèle d'autorégulation de la maladie conçoit l'adhésion comme la résultante d'un processus de résolution de problème. Ainsi, si la compréhension du problème n'est pas cohérente avec la solution (le PESL), il n'est pas surprenant que la personne adhère peu ou pas. Dans la lignée du processus de résolution de problème, l'étape suivante est d'évaluer si la solution apporte les changements escomptés. Les analyses multivariées portant sur la prédiction de l'adhésion supportent cette hypothèse. En effet, le prédicteur de loin le plus influent de l'adhésion a été la perception par le participant d'un changement global (variable ChangGlob) dans sa condition. ChangGlob a été à son tour prédit par PercMaladie en tant que prédicteur le plus influent. Ces mêmes résultats avaient été obtenus lors de l'étude préliminaire ($n = 48$ participants), la variance expliquée de l'adhésion et ChangGlob étant de 50 et 58 % respectivement (Larivière *et al.*, 2016). La présente étude, comprenant 110 participants, permet maintenant d'expliquer 76 et 50 % de la variance de l'adhésion et ChangGlob, démontrant la robustesse des résultats.

Les perceptions (ou représentations) de la maladie prédisent donc indirectement l'adhésion, puisqu'elle serait la première étape du processus de résolution problème. Contrairement à la perception du changement (ChangGlob) qui arrive plus tard dans le processus de gestion de problème, elle peut être documentée dès l'évaluation initiale du patient. Ces représentations peuvent être modifiées par l'éducation, lorsque les patients ont une compréhension partielle de leur condition et cherchent des réponses (Coutu *et al.*, 2013). L'éducation peut s'effectuer sur les représentations qui incluent la compréhension des composantes suivantes, soit les causes, l'efficacité personnelle à effectuer les comportements souhaités ou du traitement, les conséquences, la chronologie ou les

symptômes perçus associés à la lombalgie. Lorsque les patients ont des difficultés à se libérer de certaines croyances ou craintes bien ancrées, les représentations peuvent également être abordées de manière indirecte. Par exemple, utiliser les suivis en clinique lors desquels le programme est enseigné pour solliciter les appréhensions ou les objections qu'ils peuvent générer et solliciter les indicateurs d'effets que le travailleur utilise pour juger des changements globaux. Le tout vise à ce que les cliniciens proposent, expliquent ou greffent des stratégies qui ont du sens pour les patients. Ceci aidera à adopter le comportement et à en voir l'impact (Coutu *et al.*, 2013). D'ailleurs, il apparaît que les représentations peuvent changer aussi rapidement que dans les deux premières semaines d'un programme, lorsque les stratégies sont adaptées (Coutu *et al.*, 2013). Lorsqu'il est plus difficile d'accéder aux représentations ou si elles sont très fortement ancrées, les physiothérapeutes peuvent également collaborer avec d'autres disciplines (par exemple, les psychologues, l'ergothérapie et la médecine (Peek *et al.*, 2016).

Les peurs et croyances face à l'activité physique (Peurs/croyances-AP) a été introduit à T8 (modèle 9; score > 6,5 / 24) ainsi qu'à T34 dans trois modèles, mais cette fois-ci avec une interprétation opposée (score < 12 / 24), ce qui n'est pas étranger aux résultats d'autres RPC, comme détaillé plus loin. Cette variable n'a pas été retenue dans nos analyses multivariées en lien avec l'adhésion au PESL à domicile, alors le mécanisme d'action est possiblement différent. Étonnamment, le questionnaire FABQ a traditionnellement été considéré à l'étape de dérivation de différentes RPC en physiothérapie (Cai *et al.*, 2009; Cleland *et al.*, 2007; Flynn *et al.*, 2002; Hicks *et al.*, 2005; Stolze *et al.*, 2012), et parfois cela sans apparente hypothèse sous-jacente pour justifier la relation possible avec le traitement. Cela est manifeste pour les traitements passifs tels que la traction ou la manipulation vertébrale. Par contre, il est possible que ce soit pour exclure certains patients qui auraient plus besoin d'une intervention de nature psychologique que de nature biomédicale, bien que l'approche comportementale soit l'idéal en contexte de « phobie ». La sous-échelle de travail a d'ailleurs été retenue dans la RPC de succès de la traction vertébrale (seuil ≤ 21 / 42) de Cai *et al.* (2009) et dans la RPC de succès de la manipulation vertébrale (seuil < 19 / 42) de Flynn *et al.* (2002), suggérant que les patients ayant des peurs et croyances élevées au travail ne répondraient pas à ces traitements passifs, dont les mécanismes d'action seraient plus de nature biomédicale. La sous-échelle d'activité physique (Peurs/croyances-AP) a également été retenue dans la RPC de succès (seuil < 12 / 24) de manipulation vertébrale thoracique de Cleland *et al.* (2007) et dans la RPC d'échec (seuil < 9 / 24) du PESL de Hicks *et al.* (2005). Cleland *et al.* (2007) n'étaient pas en mesure d'émettre une hypothèse d'explication de leurs résultats. L'interprétation des résultats de la présente recherche (du moins pour le modèle à T8) va dans le même sens que la RPC d'échec de Hicks *et al.* (2005), si on inverse l'interprétation des résultats de Hicks (prédiction de l'échec au lieu du succès). Ainsi, le PESL aiderait surtout les patients présentant des niveaux modérés à élevés à la sous-échelle Peurs/croyances-AP à surmonter leurs peurs, ce qui en retour améliorerait les incapacités perçues, comme proposé par le modèle d'évitement de la

peur (Vlaeyen et Linton, 2000). Si ce n'est pas via l'adhésion aux exercices à domicile, cela pourrait être via l'exécution des exercices faits en clinique. Rappelons que le PESL, avec ses trois phases permettant de vaincre les défis d'ordre sensorimoteur et fonctionnel, offre de façon concomitante une exposition graduelle permettant de confronter les peurs, ce qui représente des ingrédients gagnants pour ce type de clientèle. Par contre, tout comme la RPC de succès pour les manipulations vertébrales thoraciques (Cleland *et al.*, 2007), les résultats opposés observés à T34 (score < 12 / 24) suggèrent que les participants doivent déjà ne pas avoir de peurs trop élevées au départ pour obtenir un succès. Si ces peurs sont trop grandes, il faut envisager de réduire l'intensité du PESL pour l'aborder encore plus graduellement. Le rôle des Peurs/croyances-AP dans les différentes RPC en physiothérapie reste définitivement à éclaircir.

5.1.3 Modèles comportant des variables de classe A, B et C

Les variables de classe C sont essentiellement des variables mesurant l'amplitude de mouvement à différentes articulations. Or, aucune étude n'a établi un lien entre la mobilité et l'instabilité lombaires, à part la mobilité entre deux vertèbres lombaires bien entendu. Ces variables sont donc non spécifiques, ce qui signifie qu'elles pourraient influencer l'issue de tout programme d'exercices. Ces variables informent sur le pronostic du patient et pourraient donc être considérées dans les RPC « pronostiques », c.-à-d. sans égard au traitement suivi, mais ne peuvent que discréditer les RPC « prescriptives » comme celles qui ont été dérivées dans la présente étude (Haskins *et al.*, 2012). En effet, considérant le devis de la présente étude (étude de cohorte), c.-à-d. sans la considération d'un groupe contrôle permettant d'exclure les changements cliniques non spécifiques au PESL, il a été jugé dès le départ que les variables de classe C ne pouvaient pas faire partie des RPC finales sélectionnées parmi tous les modèles étudiés. Cette stratégie aiderait à réduire la probabilité que ces RPC finales ne soient pas spécifiques au PESL.

L'intérêt d'avoir produit des modèles avec des variables de classe C est de mettre en évidence qu'elles peuvent avoir une valeur prédictive, même après avoir inclus les variables de classe A et B dans le modèle, ce qui a été le cas pour la prédiction du succès à T34 seulement. Dans une perspective de dérivation d'une RPC pronostique, cette information revêt donc de la valeur.

Les modèles 14 et 16 de prédiction du succès à T34 incluaient deux variables provenant du test passif d'élévation jambe tendue, soit l'étendue de mouvement moyen des deux côtés (EJT-Pas-Moy-EDM < 93° ; modèle 14) et l'étendue de mouvement minimal (entre les deux côtés) lorsque la douleur apparaît (EJT-Pas-Doul-EDM < 68° ; modèle 16). Ces deux variables proviennent du même test et sont hautement corrélées avant leur dichotomisation (r de Pearson : 0,83 ; $P < 0,001$), révélant qu'elles capturent essentiellement la même information. La première variable avait été retenue dans la RPC de succès de Hicks *et al.* (2005), soit EJT-Pas-Moy-EDM > 91°, ce qui apparaît très

ressemblant (91° vs 93°), mais dans la direction opposée. Une diminution de cette variable est associée à la présence d'une radiculopathie et à un pronostic défavorable, ce qui rend les résultats des modèles 14 et 16 contre-intuitifs en ce qui a trait à l'interprétation de ces variables. Cependant, ce test a une sensibilité élevée à la radiculopathie liée à une hernie discale, ceci combiné à une faible spécificité (Deyo *et al.*, 1992). Étant donné que nos participants ne présentent pas de problèmes neurologiques (critère d'exclusion), ces résultats ne devraient pas être associés à une radiculopathie. Ce test pourrait simplement être une indication d'une flexibilité réduite des ischio-jambiers et/ou d'une mobilité réduite du nerf sciatique. Il a également été associé à une diminution du flux sanguin vers les racines nerveuses (Kobayashi *et al.*, 2003), ce qui pourrait, en théorie, être lié à une irritation des racines nerveuses adjacentes à un segment lombaire instable (Shacklock, 1995).

5.1.4 Quelle RPC choisir ? Pour prédire le succès à la fin du traitement (T8) ou au suivi de six mois après la fin du traitement (T34) ?

Une revue critique de la littérature portant sur les RPC développées pour les interventions en physiothérapie souligne que toutes ces RPC ne prédisent le succès qu'à la fin du traitement, alors qu'il serait très informatif de prédire ce succès après une période de suivi d'au moins six mois après le traitement (Beneciuk *et al.*, 2009). À la connaissance des auteurs, la dérivation de la RPC de succès à T34 (six mois de suivi) est une première en physiothérapie. Idéalement, les deux RPC de succès (à T8 et T34) devraient être développées en passant aux étapes 2 (validation) et 3 (étude d'impact) du développement d'une RPC (Childs et Cleland, 2006). Ces étapes de développement requérant énormément de ressources, il est beaucoup plus réaliste de ne sélectionner qu'une RPC pour la suite.

Quelques arguments supportent le choix de la RPC à T8. Lorsqu'au moins deux (RPC à T8) ou trois (RPC à T34) variables sont présentes, bien que le rapport de vraisemblance positif (LR+) puisse être équivalent (de 17,0 à 17,9), la RPC à T8 offre une hausse de probabilité de succès plus importante (de 49 à 96 %, soit une hausse de 47 %) que la RPC à T34 (de 53 à 85 %, soit une hausse de 32 %). En termes de facilité d'utilisation par les thérapeutes, bien que la RPC à T34 ne comporte que trois tests cliniques au lieu de quatre pour la RPC à T8, un de ces trois tests est plus pénible pour le patient, soit le test d'endurance des muscles abdominaux (TEM-abdominaux).

Chez les 100 patients qui ont fait le suivi de 6 mois, les effets du PESL ont été très positifs à T8 et se sont maintenus à T34, ceci autant pour les perceptions d'incapacité (Oswestry : tailles d'effets de -1,20 et -1,30 à T8 et T34 respectivement) que pour l'intensité de la douleur (tailles d'effets de -1,70 et -1,35 à T8 et T34 respectivement). Ces résultats sont en accord avec une méta-analyse (Niederer et Mueller, 2020) indiquant que les PESL maintiennent ces effets post-traitement (incapacités, douleur) mieux qu'aucun autre comparateur (traitements passifs, inactifs ou autres types d'exercices). Il est donc

raisonnable de penser que l'utilisation de la RPC à T8 ait une certaine valeur prédictive sur les résultats à T34.

5.2 Objectif 2 – Étude des mécanismes d'action

Très peu d'essais cliniques ont prévu l'évaluation des mécanismes d'action sous-jacents à différents programmes d'exercices chez des patients souffrant d'une lombalgie chronique, comme conclue une récente revue sur le sujet (Wun *et al.*, 2020). Les comparaisons avec d'autres études sont donc rarement possibles. Bien que le devis de la présente étude ne soit pas un essai clinique randomisé contrôlé, il permet, par la comparaison entre les sous-groupes succès et échec, d'émettre des hypothèses sur les mécanismes sous-jacents à un PESL.

5.2.1 Évaluation physique (mesures PHT)

Les interactions SOUS-GROUPE (GE, GS) $\times$ TEMPS (T0, T8) des ANOVA ont été significatives ($P \leq 0,05$) pour plusieurs tests physiques demandant une performance maximale des participants (variables TPP et TEM), révélant une plus grande amélioration chez le GS comparativement au GE (**tableau 15, figure 7**). Ces résultats supportent l'hypothèse que le PESL améliore la condition musculaire, du moins en termes d'endurance (variables TEM) et possiblement en termes de coordination et puissance musculaires (variable TPP), et que ces améliorations sont associées au succès du traitement. Plusieurs études démontrent l'effet positif d'un PESL sur l'endurance des muscles du tronc (Gopika et Dinakaran, 2018; Hoppes *et al.*, 2016; Javadian *et al.*, 2012; Sandrey et Mitzel, 2013). Étant donné que ces mesures de performance pourraient être influencées par des variables psychologiques associées à la douleur (Dramatisation/douleur, Peurs/croyances-AP), des analyses corrélationnelles complémentaires ont été réalisées pour évaluer cette possibilité. Or, des corrélations significatives, mais faibles ont été obtenues entre Δ OSW ($\Delta = T8-T0$) et les variables Δ TPP ($r = 0,21$ à $0,35$) et Δ TEM-dorsaux ($r = -0,31$) et ces corrélations n'ont que légèrement été diminuées (de $0,03$ à $0,10$ point) lorsqu'elles étaient ajustées avec Δ Dramatisation/douleur ou Δ Peurs/croyances-AP, ce qui rejette cette possibilité. Ces résultats positifs sont en contradiction avec la conclusion d'une revue systématique qui a étudié la corrélation entre les changements observés avec différentes mesures de performance (force, mobilité, endurance) et les changements cliniques observés (douleur, incapacités perçues) au cours de différents programmes d'exercices (Steiger *et al.*, 2012). Leur conclusion générale était que les résultats n'appuient pas la notion que les effets cliniques de la thérapie par l'exercice soient directement attribuables aux changements dans le système musculosquelettique. Or, la force n'a pas directement été mesurée dans la présente étude et aucune étude n'avait fait la relation entre l'endurance et les incapacités perçues dans la revue de Steiger *et al.* (2012). Il est possible que notre approche de contraste entre deux sous-groupes très différents (en excluant le sous-groupe « amélioration clinique » se situant entre les deux) ait permis de mieux faire

ressortir les résultats. Par contre, il est vrai que l'étude de nos mesures de mobilité mène à la même conclusion que cette revue, comme discuté dans le paragraphe suivant.

L'absence d'interaction significative pour les mesures d'étendue de mouvement (**tableau 15**) démontre que bien qu'une amélioration de la mobilité soit observée au cours du PESL (flexion et rotation interne de la hanche, flexion latérale lombaire), cette amélioration ne serait pas associée au succès du PESL. Ces résultats supportent notre classement des variables (classe C) lors du développement de la RPC, car à la connaissance des auteurs, il y a absence de résultat établissant une relation entre la mobilité et l'instabilité lombaires. D'ailleurs, le PESL ne comportait pas d'exercices de mobilité.

L'examen physique par les différents tests de provocation de la douleur a également mené à des résultats supportant l'idée que le succès du PESL passerait par une réduction plus marquée des symptômes lors de la plupart des tests de déficiences du contrôle sensorimoteur, de même que lors du test d'instabilité lombaire (Instabilité-Ext) et du test actif d'élévation jambe tendue (**tableau 16**). La présence de mouvements aberrants serait également réduite chez le GS seulement. L'ensemble de ces résultats permet donc de s'avancer avec une certaine assurance sur la présence d'une relation entre la réparation ou guérison des tissus et les incapacités. Selon la théorie sous-jacente au PESL, la réparation des tissus s'expliquerait par un meilleur contrôle du mouvement relatif entre les vertèbres lombaires, évitant l'exacerbation des lésions à la source de la lombalgie et laisserait ainsi le temps nécessaire à cette réparation.

5.2.2 Évaluation psychologique (mesures PSY)

Les interactions SOUS-GROUPE (GE, GS) × TEMPS (T0, T8) des ANOVA ont été significatives ($P \leq 0,05$) pour quelques variables de nature psychologique (Peurs/croyances-AP, Dramatisation/douleur, PercMaladie, SEPE), de même que pour l'outil de dépistage STarT Back et pour l'intensité de la douleur, révélant une plus grande amélioration chez le GS comparativement au GE (**tableau 17, figure 8**). Alors que les sous-groupes étaient équivalents à T0 dans tous ces cas, un écart se creusait entre eux d'un temps de mesure à l'autre (T4 et T8) et se stabilisait au suivi de 6 mois (T34), avec des tailles d'effet dépassant les seuils de 0,5 (moyenne) et 0,8 (forte) dans plusieurs cas. Compte tenu des résultats pour les variables Peurs/croyances-AP, Dramatisation/douleur et PDI_{tot}, il n'est pas étonnant que l'outil STarT Back ait révélé la même tendance, car cet outil est plus ou moins une mesure agrégat de ces différents concepts. Il est important de rappeler que ces effets ne sont probablement pas spécifiques à un PESL, c.-à-d. qu'ils sont possibles pour tout type de programme d'exercices.

5.2.2.1 Effets sur les peurs et croyances face à l'activité physique (Peurs/croyances-AP) et de la dramatisation de leurs douleurs (Dramatisation/douleur).

Il apparaît donc que les participants qui vivront un succès avec le PESL ont une diminution plus importante de leurs peurs et croyances face à l'activité physique (Peurs/croyances-AP) et de la dramatisation de leurs douleurs (Dramatisation/douleur). Ceci supporte l'idée que le PESL permet une exposition graduelle à l'activité qui est favorable à la réduction des craintes liées à la douleur et au mouvement, comme le propose le modèle d'évitement de la peur (Vlaeyen et Linton, 2000) et comme démontré pour différents programmes d'exercices (Mannion *et al.*, 1999; Marshall *et al.*, 2013; Smeets *et al.*, 2006; Woby *et al.*, 2004). Par contre, les résultats de la présente étude démontrent que certains patients (GS) profitent plus de cette exposition graduelle que d'autres (GE), ce qui peut expliquer leur succès, du moins en partie. Ces deux sous-groupes étaient égaux au départ pour ces variables, mais un modèle de prédiction de succès a tout de même retenu Peurs/croyances-AP à T8 (modèle 9; score > 6,5 / 24), ce qui donne une indication au-delà de quel seuil le PESL peut devenir bénéfique. Par contre, trois modèles à T34 proposent un seuil < 12 / 24, suggérant que ce succès ne peut se maintenir à plus long terme si ces peurs sont trop élevées. Il faut peut-être prévoir une progression plus graduelle chez ces patients.

Avec un programme inspiré de l'approche australienne (Lee, H. S. *et al.*, 2016; Unsgaard-Tondel *et al.*, 2010) aucun effet sur Peurs/croyances-AP n'avait été détecté. Les effets observés ici et dans une autre étude (Alrwaily *et al.*, 2019) peuvent possiblement être attribuables au développement de l'endurance musculaire (20 à 30 répétitions) qui caractérise l'approche de McGill. Rappelons que le présent PESL combine les deux écoles de pensées. D'ailleurs, rappelons que Peurs/croyances-AP a été sélectionné dans la présente RPC de succès (Peurs/croyances-AP > 6/24) et la RPC d'échec (Peurs/croyances-AP < 9/24) de Hicks *et al.* (2005), les deux PESL étudiés (Hicks *et al.*, 2005 et la présente étude) comportant une composante de développement de l'endurance musculaire. Cet élément n'est pas négligeable, car les exercices de stabilisation lombaire ne contiennent pas d'activités physiques vigoureuses ni de mouvements autres que l'adoption de postures statiques de façon prolongée, notamment en contraignant la région douloureuse à la source de ces craintes.

Il apparaît que le chargement imposé sur les structures lombaires (incluant les muscles) durant les différentes postures, lorsque ces dernières sont maintenues de sorte à développer l'endurance musculaire, serait suffisant pour réduire ces craintes. Cette hypothèse est soutenue par les effets positifs aussi observés sur Peurs/croyances-AP et Dramatisation/douleur suite à un programme d'exercices de Pilates, soit une forme d'exercices très comparable au présent PESL (Marshall *et al.*, 2013). Des études avec suivi à plus long terme permettraient de vérifier si cette expérience de succès est suffisante pour que ces participants, après avoir éventuellement arrêté les exercices à

domicile, puissent reprendre ces exercices de leur propre chef, si ces craintes ne représentent plus un frein. Ceci fait appel à l'adhésion aux exercices, comme discuté dans la prochaine section.

5.2.2.2 Effets sur les perceptions de la maladie (PercMaladie) et le sentiment d'efficacité personnelle à faire les exercices (SEPE).

Les effets sur les variables PercMaladie et SEPE semblent liés à l'adhésion aux exercices. En effet, la poursuite du PESL, qui est mesurée dans notre étude par l'adhésion à T8 et T34, fut clairement associée à une perception de changement dans la condition de santé dans le temps (ChangGlob), tout comme à l'incapacité (OSW), mais également au sentiment d'efficacité personnelle à faire les exercices (SEPE). Chez le GE, une baisse graduelle de SEPE a été observée dès T4 et T8 pour ensuite démontrer une chute brutale à T34 (**figure 8**). Chez le GS, SEPE s'est maintenu jusqu'à T8, pour ensuite observer une diminution à T34, mais beaucoup moindre (**figure 8**). Il s'avère tout de même que cette chute se produit au moment où il n'y a plus de suivis en clinique et que les patients doivent poursuivre par eux-mêmes.

Dans une perspective clinique, ce qui est intéressant d'observer est que la perception de changement (ChangGlob), qui est la principale variable prédictrice de l'adhésion, est fortement liée au SEPE et à la PercMaladie. Ces deux variables sont des concepts clés du modèle d'autorégulation de la maladie. Ainsi, le résultat soutient la théorie sous-jacente au modèle d'autorégulation de la maladie. Les résultats de l'ANOVA sont très clairs, démontrant une diminution beaucoup plus marquée de PercMaladie (effet positif) pour le GS que pour le GE. Malheureusement, ChangGlob et PercMaladie ne sont généralement pas évaluées dans ce type de recherche portant sur les effets de programmes d'exercices. En ce sens, les résultats observés dans la présente étude sont tout à fait novateurs pour le domaine.

5.2.3 Évaluation neuromusculaire (mesures NRM)

Pour les mesures neuromusculaires, les ANOVA se sont révélées peu révélatrices en matière d'interactions SOUS-GROUPE × TEMPS et d'effets TEMPS. En effet, les quelques cas qui ont atteint le niveau de signification statistique n'ont révélé que des effets de faible amplitude, voire négligeable. L'ensemble de ces résultats NRM suggère donc que le PESL n'induirait pas de modifications neuromusculaires qui à leur tour seraient responsables du succès de ce traitement. Les résultats préliminaires (Larivière *et al.*, 2016), basés sur des analyses corrélationnelles, ont mené à quelques corrélations significatives ($P < 0,05$), mais d'amplitude modérée ($r = 0,33$ à $0,49$), ce qui suggérerait que de tels mécanismes d'action pouvaient être présents. Ces corrélations ont été reprises avec l'échantillon entier de patients ($n = 77$), menant à des résultats similaires, mais d'amplitude un peu plus faible ($r = 0,23$ à $0,37$), mais n'ont pas été présentées ici. Par contre, les analyses actuelles, basées sur la comparaison entre sous-groupes extrêmes

dans le temps (GC : groupe contrôle; GS : groupe succès; GE : groupe échec), ne confirment pas les tendances qui avaient été observées. Bien qu'il ne s'agisse pas de très grands groupes (entre 30 et 40 participants), ils sont jugés assez importants dans cette littérature, car les mesures NRM sont coûteuses. Ces sous-groupes sont jugés suffisants pour conclure que si des effets réels (interactions) n'ont pas été détectés (erreur de type 2), c'est parce que ces effets sont relativement petits, à la lumière des petites tailles d'effet associées aux effets TEMPS relevés. En d'autres mots, ces effets non détectés seraient négligeables sur le plan clinique.

Les études portant sur l'effet d'un PESL sur les paramètres neuromusculaires sont peu nombreuses (Bruno, 2014) et plusieurs chercheurs s'entendent sur le fait que les effets sont beaucoup déterminés par le choix des patients recrutés. En effet, étudier un groupe de patients hétérogènes comporte d'importantes limites, car tous ces patients n'auront pas les déficiences visées par un PESL. En d'autres mots, il est évident que des effets plus marqués sont attendus chez des patients démontrant des déficiences en termes d'instabilité lombaire. La stratégie la plus utilisée (Javadian *et al.*, 2012; Park et Lee, 2021; Teyhen *et al.*, 2007) est de ne recruter que des patients démontrant des signes d'instabilité lombaire, tels que les patients qui répondent aux critères de la RPC de succès de Hicks *et al.* (2005). Une autre consiste à utiliser un système de classification aussi basé sur la probabilité de succès à différents traitements (treatment-based classification (Delitto *et al.*, 1995), lequel repose sur trois tests cliniques, soient un score de Beighton > 4/9, la présence de mouvement aberrant et le « posterior shear test ». Il est intéressant que les RPC de succès dérivées dans la présente étude confirment la pertinence des deux premiers tests, le dernier n'ayant pas été évalué dans la présente étude, faute d'une fidélité interjuges suffisante.

L'objectif principal de la présente étude étant justement de dériver une RPC de succès, il n'était pas possible de sélectionner les patients au départ. Par contre, dans la présente étude, il serait attendu que des interactions SOUS-GROUPE × TEMPS significatives s'expliquent par une amélioration plus importante du GS, justement parce que ce dernier démontrerait certaines déficiences neuromusculaires au départ, contrairement au GE. Bien entendu, on ne peut s'attendre à ce que le GS ait des mesures anormales sur tous les paramètres neuromusculaires évalués dans la présente étude. Par contre, les différences significatives observées entre les trois sous-groupes (GC vs GS vs GE) ne supportent pas cette hypothèse. Les résultats détaillés des mesures neuromusculaires seront publiés dans des articles scientifiques et des analyses plus poussées devront être menées pour tenter de mieux comprendre ce qui se produit et ainsi émettre de nouvelles hypothèses.

5.3 Forces et limites

La force principale de cette étude réside dans le fait que les sous-groupes cliniques (succès, échec) ont été définis avec des critères permettant de bien les distinguer en éliminant le sous-groupe « amélioration clinique » situé entre les deux. Cette approche, proposée par Hicks *et al.* (2005), est plus rigoureuse que plusieurs autres études de dérivation de RPC qui définissent les sous-groupes succès et échec en séparant tout simplement le groupe en deux. Nous croyons que cela permet de sélectionner des prédicteurs qui ont un plus grand lien avec le traitement spécifique qui est proposé. En effet, le devis utilisé (essai clinique sans groupe contrôle) a le potentiel d'ouvrir la porte à des prédicteurs qui sont non spécifiques aux effets du programme d'exercice. Il est alors important de minimiser les risques que cela se produise. C'est d'ailleurs aussi la raison pour laquelle les modèles de régression étudiés comportaient trois niveaux hiérarchiques (incluant les variables de classe A, B et C) qui tenaient compte de la force de la relation théorique entre chaque prédicteur et les effets attendus d'un PESL. Cette caractéristique distingue la présente étude de dérivation des études précédentes. Comme discuté plus haut, cela permet des interprétations cliniques plus nuancées, ce qui a d'ailleurs mené à une RPC de succès ne comportant que des variables de classe A qui ont un lien théorique plus fort avec le concept de stabilité lombaire. Une autre force de cette étude est la considération d'un large spectre de prédicteurs potentiels du succès ou de l'échec du traitement, qu'ils soient de nature physique ou psychologique. Finalement, l'étude des mécanismes de nature physique, psychologique et neuromusculaire qui ont le potentiel d'expliquer le succès du PESL permet d'apporter un éclairage que peu d'études ont tenté de faire. La batterie de six tests neuromusculaires rend la présente étude unique.

La limite principale de cette étude est l'utilisation d'un devis expérimental sans groupe de patients n'ayant pas suivi le programme d'exercice (vrai groupe contrôle). Il faut donc être très prudent sur la façon d'interpréter les résultats, car les effets observés peuvent, du moins en partie, s'expliquer par une récupération naturelle et spontanée des patients, sans contribution positive du programme d'exercice. Cet avertissement vaut autant pour l'interprétation de la RPC de succès que pour l'étude des mécanismes d'action sous-jacents au PESL. Par contre, le fait que plus de 97 % (107/110) des patients étaient en phase chronique (lombalgie d'une durée de plus de trois mois) et que plus de 85 % (94/110) des patients avaient une lombalgie d'une durée de plus d'un an, une récupération spontanée menant aux fortes tailles d'effet observées ici est improbable (d de Cohen de -1,24 et -1,70 pour OSW et Int-Douleur à T8, respectivement).

Les tests d'évaluation des déficiences du contrôle sensorimoteur ont été limités à leurs manifestations en termes de symptômes, car les compétences nécessaires pour l'évaluation des signes (alignement, mouvements) apparaissent beaucoup plus difficiles à acquérir (Van Dillen *et al.*, 1998), ce qui rendrait l'application d'une éventuelle RPC plus complexe pour la majorité des cliniciens. Ce choix peut apparaître comme une force pour certains et comme une limite pour d'autres. En effet, cela a eu pour effet d'éliminer les

tests de dissociation thoracolombaire proposés par Luomajoki *et al.* (2007), dont certains ont démontré une bonne fidélité interjuges ($\kappa > 0,6$). Les tests de performance physique, notamment ceux basés sur la vitesse d'exécution (TPP-assis, TPP-flexions et TPP-rouler), peuvent être critiqués en ce qui a trait au lien qui a été accordé avec l'instabilité lombaire (variables de classe A). Il a été jugé que l'exécution de mouvements rapides nécessite une bonne coordination musculaire pour assurer une rigidité et une stabilité lombaire adéquate. Ce sont des études de validation avec une mesure directe de la stabilité lombaire, notamment par imagerie médicale des mouvements intervertébraux en condition dynamique (Dombrowski *et al.*, 2018; Wu *et al.*, 2014), qui pourront aider à supporter ou infirmer ces hypothèses.

Une autre limite concerne l'étude des mécanismes d'action, car elle comportait un nombre très important de variables, ce qui augmente les risques d'erreurs statistiques de type 1 (détection d'effets significatifs par hasard). Cette approche exploratoire permet toutefois de se servir de la tendance d'un ensemble de variables pour se prononcer sur l'effet global du PESL, car il serait improbable qu'un grand nombre d'effets significatifs aient atteint le niveau de signification statistique par hasard, ce risque étant estimé à 5 % ($\alpha = 0,05$). Le nombre important d'effets détectés pour les variables PHT et pour les variables PSY suggère donc la présence d'effets réels, pris dans leur ensemble et non pour chaque variable prise individuellement. Cela peut être attribuable à une sélection de variables basée sur des modèles théoriques pertinents, permettant de limiter la détection d'effets par le fruit du hasard. Il reste tout de même que cette approche est exploratoire, permettant d'émettre des hypothèses de mécanismes d'action, mais pas de confirmer leur présence. Ces hypothèses demandent à être évaluées plus formellement avec le devis approprié, dans de futurs essais cliniques randomisés contrôlés. Une autre limite importante concerne la RPC de succès au suivi de six mois, car les patients étaient libres d'avoir recours à tout type de soin. Il aurait été non éthique de procéder autrement. Par contre, bien que cela corresponde à la réalité, il devient plus difficile d'attribuer le succès uniquement au PESL.

Finalement, les patients ont en très grande partie été recrutés par le biais d'annonces dans les journaux (très peu en clinique). Il est donc possible que l'état clinique de ces patients, ou leur comportement face à ce programme d'exercice ne reflètent pas l'état clinique ou le comportement de patients uniquement recrutés dans les cliniques. Par exemple, comme détaillé au tout début de la discussion, la sévérité de la condition des participants (douleur, incapacité) était probablement plus élevée que ce qui est rencontré en clinique, car un critère de sévérité ($\text{OSW} > 12\%$) a été utilisé au recrutement. Aussi, peu de participants étaient dans la phase subaiguë. Le devis de l'étude (approche per protocole) imposait également que les participants adhèrent au programme clinique. Il est donc possible que les patients recrutés à l'extérieur de la clinique aient été plus « volontaires » et donc moins réticents à débiter un programme d'exercice, ce qui pourrait influencer les relations entre certaines mesures de nature psychologique et les mesures de résultats cliniques.

CONCLUSION

La présente recherche a permis de compléter avec succès l'étape de dérivation de RPC de succès d'un PESL, ceci autant pour le succès observé à la fin du traitement (ici à T8) que pour le succès observé après une période de suivi de six mois (ici T34). La RPC de succès à T34 représente le premier exemple de RPC pour prédire le suivi en post-traitement dans le domaine de la physiothérapie. Ces résultats justifient la poursuite du développement de tels outils permettant de faire un meilleur arrimage entre certains patients et ce programme d'exercices en particulier afin d'en maximiser les bénéfices. En effet, il est important de rappeler qu'une RPC ne peut en principe être diffusée et utilisée avec confiance qu'après la réalisation d'études de validation et d'analyse d'impact (Childs et Cleland, 2006).

L'étude des mécanismes d'action suggère que ce programme d'exercice engage des mécanismes d'action bénéfiques pour les patients, ceci autant sur le plan physique que psychologique. Le PESL étudié dans la présente recherche est basé sur une approche hybride combinant l'approche australienne, qui se concentre sur le contrôle isolé des muscles profonds du tronc, suivie par l'approche de McGill pour rehausser l'endurance musculaire. En accord avec les résultats de la présente recherche, ces caractéristiques permettraient d'abord une exposition graduelle à l'activité qui est favorable à la réduction des craintes liées à la douleur et au mouvement, comme le propose le modèle d'évitement de la peur, et ensuite au développement de l'endurance des muscles du tronc. Par contre, les mesures neuromusculaires réalisées en laboratoire n'ont pas révélé de mécanismes associés à un meilleur contrôle de la coordination des muscles du tronc chez les patients qui ont eu un succès thérapeutique, comparativement aux patients qui ont eu un échec. Finalement, l'étude plus spécifique de l'adhésion au programme d'exercices à domicile suggère que certains des mécanismes de nature psychologique en question joueraient un rôle dans cette adhésion, ce qui a son tour aurait un impact sur le succès du PESL.

BIBLIOGRAPHIE

- Adkins, D. L., Boychuk, J., Remple, M. S. et Kleim, J. A. (2006). Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. *Journal of Applied Physiology*, 101(6), 1776-1782. <https://doi.org/10.1152/japplphysiol.00515.2006>
- Airaksinen, O., Brox, J. I., Cedraschi, C., Hildebrandt, J., Klaber-Moffett, J., Kovacs, F., . . . Zanolli, G. (2004). Chapter 4: European guidelines for the management of chronic non-specific low back pain. *European Spine Journal*, 15(S2), s192-s300. <https://doi.org/10.1007/s00586-006-1072-1>
- Alqarni, A. M., Schneiders, A. G. et Hendrick, P. A. (2011). Clinical tests to diagnose lumbar segmental instability: A systematic review. *Journal of Orthopaedic & Sports Physical Therapy*, 41(3), 130-140. <https://www.jospt.org/doi/10.2519/jospt.2011.3457>
- Alrwaily, M., Schneider, M., Sowa, G., Timko, M., Whitney, S. L. et Delitto, A. (2019). Stabilization exercises combined with neuromuscular electrical stimulation for patients with chronic low back pain: A randomized controlled trial. *Brazilian Journal of Physical Therapy*, 23(6), 506-515. <https://doi.org/10.1016/j.bjpt.2018.10.003>
- Alyazedi, F. M., Lohman, E. B., Wesley Swen, R. et Bahjri, K. (2015). The inter-rater reliability of clinical tests that best predict the subclassification of lumbar segmental instability: Structural, functional and combined instability. *Journal of Manual & Manipulative Therapy*, 23(4), 197-204.
- American Physical Therapy Association. (2000). Clinical research agenda for physical therapy. *Physical Therapy*, 80(5), 499-513. <https://doi.org/10.1093/ptj/80.5.499>
- Arjmand, N. et Shirazi-Adl, A. (2006). Model and in vivo studies on human trunk load partitioning and stability in isometric forward flexions. *Journal of Biomechanics*, 39(3), 510-521. <https://doi.org/10.1016/j.jbiomech.2004.11.030>
- Asgari, M., Sanjari, M. A., Mokhtarinia, H. R., Moeini Sedeh, S., Khalaf, K. et Parnianpour, M. (2015). The effects of movement speed on kinematic variability and dynamic stability of the trunk in healthy individuals and low back pain patients. *Clinical Biomechanics*, 30(7), 682-688. <https://doi.org/10.1016/j.clinbiomech.2015.05.005>
- Baecke, J. A. H., Burema, J. et 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(5), 936-942.
- Barker, P. J., Guggenheimer, K. T., Grkovic, I., Briggs, C. A., Jones, D. C., Thomas, C. D. et Hodges, P. W. (2006). Effects of tensioning the lumbar fasciae on segmental stiffness during flexion and extension: Young Investigator Award winner. *Spine*, 31(4), 397-405. <https://doi.org/10.1097/01.brs.0000195869.18844.56>

- Barr, K. P., Griggs, M. et Cadby, T. (2007). Lumbar stabilization: A review of core concepts and current literature, part 2. *American Journal of Physical Medicine & Rehabilitation*, 86(1), 72-80. <https://doi.org/10.1097/01.phm.0000250566.44629.a0>
- Beinart, N. A., Goodchild, C. E., Weinman, J. A., Ayis, S. et Godfrey, E. L. (2013). Individual and intervention-related factors associated with adherence to home exercise in chronic low back pain: A systematic review. *Spine Journal*, 13(12), 1940-1950.
- Beneciuk, J. M., Bishop, M. D. et George, S. Z. (2009). Clinical prediction rules for physical therapy interventions: A systematic review. *Physical Therapy*, 89(2), 114-124. <https://doi.org/10.2522/ptj.20080239>
- Benish, W. A. (2020). A review of the application of information theory to clinical diagnostic testing. *Entropy*, 22(1). <https://doi.org/10.3390/e22010097>
- Bertilson, B. C., Bring, J., Sjoblom, A., Sundell, K. et Strender, L. E. (2006). Inter-examiner reliability in the assessment of low back pain (LBP) using the Kirkaldy-Willis classification (KWC). *European Spine Journal*, 15(11), 1695-1703.
- Biely, S. A., Silfies, S. P., Smith, S. S. et Hicks, G. E. (2014). Clinical observation of standing trunk movements: What do the aberrant movement patterns tell us? *Journal of Orthopaedic & Sports Physical Therapy*, 44(4), 262-272. <https://www.jospt.org/doi/10.2519/jospt.2014.4988>
- Biering-Sorensen, F. (1984). Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine*, 9(2), 106-119.
- Bigos, S. J., Holland, J., Holland, C., Webster, J. S., Battie, M. et Malmgren, J. A. (2009). High-quality controlled trials on preventing episodes of back problems: Systematic literature review in working-age adults. *Spine Journal*, 9(2), 147-168. <https://doi.org/10.1016/j.spinee.2008.11.001>
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16(3), 252-260.
- Boucher, J.-A., Preuss, R., Henry, S. M., Nugen, M. et Larivière, C. (2018). Trunk postural adjustments: Medium-term reliability and correlation with changes of clinical outcomes following an 8-week lumbar stabilization exercise program. *Journal of Electromyography and Kinesiology*, 41, 66-76. <https://doi.org/10.1016/j.jelekin.2018.04.006>
- Boucher, J.-A., Preuss, R., Henry, S. M., Dumas, J. P. et Larivière, C. (2016). The effects of an 8-week stabilization exercise program on lumbar movement sense in patients with low back pain. *BMC Musculoskeletal Disorder*, 17(1), article 23. <https://doi.org/10.1186/s12891-016-0875-4>
- Boucher, J.-A., Roy, N., Preuss, R. et Larivière, C. (2017). The effect of two lumbar belt designs on trunk repositioning sense in people with and without low back pain. *Annals of Physical and Rehabilitation Medicine*, 60(5), 306-311. <https://doi.org/10.1016/j.rehab.2017.03.002>

- Boudreau, S. A., Farina, D. et Falla, D. (2010). The role of motor learning and neuroplasticity in designing rehabilitation approaches for musculoskeletal pain disorders. *Manual Therapy*, 15(5), 410-414. <https://doi.org/10.1016/j.math.2010.05.008>
- Brawley, L. R. et Culos-Reed, S. N. (2000). Studying adherence to therapeutic regimens: Overview, theories, recommendations. *Controlled Clinical Trials*, 21(5, Supplement 1), 156S-163S. [http://dx.doi.org/10.1016/S0197-2456\(00\)00073-8](http://dx.doi.org/10.1016/S0197-2456(00)00073-8)
- Brennan, G. P., Fritz, J. M., Hunter, S. J., Thackeray, A., Delitto, A. et Erhard, R. E. (2006). Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: Results of a randomized clinical trial. *Spine*, 31(6), 623-631. <https://doi.org/10.1097/01.brs.0000202807.72292.a8>
- Broadbent, E., Petrie, K. J., Main, J. et Weinman, J. (2006). The brief illness perception questionnaire. *Journal of Psychosomatic Research*, 60(6), 631-637. <https://doi.org/10.1016/j.jpsychores.2005.10.020>
- Brooks, C., Kennedy, S. et Marshall, P. W. (2012). Specific trunk and general exercise elicit similar changes in anticipatory postural adjustments in patients with chronic low back pain: A randomized controlled trial. *Spine*, 37(25), E1543-E1550. <https://doi.org/10.1097/BRS.0b013e31826feac0>
- Bruno, P. (2014). The use of "stabilization exercises" to affect neuromuscular control in the lumbopelvic region: A narrative review. *The Journal of the Canadian Chiropractic Association*, 58(2), 119-130.
- Burton, A. K., Waddell, G., Tillotson, K. M. et Summerton, N. (1999). Information and advice to patients with back pain can have a positive effect: A randomized controlled trial of a novel educational booklet in primary care. *Spine*, 24(23), 2484-2491.
- Cai, C., Pua, Y. H. et Lim, K. C. (2009). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with mechanical lumbar traction. *European Spine Journal*, 18(4), 554-561. <https://doi.org/10.1007/s00586-009-0909-9>
- Centre d'action en prévention et réadaptation de l'incapacité au travail. (2007). *Le guide du dos*. CAPRIT.
- Childs, J. D. et Cleland, J. A. (2006). Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Physical Therapy*, 86(1), 122-131. <https://doi.org/10.1093/ptj/86.1.122>
- Cholewicki, J. et McGill, S. M. (1996). Mechanical stability of the in vivo lumbar spine: Implications for injury and chronic low back pain. *Clinical Biomechanics*, 11(1), 1-15. [http://www.clinbiomech.com/article/0268-0033\(95\)00035-6/abstract](http://www.clinbiomech.com/article/0268-0033(95)00035-6/abstract)
- Cholewicki, J., Panjabi, M. M. et Khachatryan, A. (1997). Stabilizing function of trunk flexor-extensor muscles around a neutral spine posture. *Spine*, 22(19), 2207-2212.

- Claus, A. P., Hides, J. A., Moseley, G. L. et Hodges, P. W. (2009). Different ways to balance the spine: Subtle changes in sagittal spinal curves affect regional muscle activity. *Spine*, 34(6), E208-E214. <https://doi.org/10.1097/BRS.0b013e3181908ead>
- Cleeland, C. S. et Ryan, K. M. (1994). Pain assessment: Global use of the Brief Pain Inventory. *Annals, Academy of Medicine, Singapore*, 23(2), 129-138.
- Cleland, J. A., Childs, J. D., Fritz, J. M., Whitman, J. M. et 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. *Physical Therapy*, 87(1), 9-23. <https://doi.org/10.2522/ptj.20060155>
- Commission de la santé et de la sécurité du travail. (2012). *Statistiques sur les affections vertébrales - 2008-2011*. CSST.
- Cook, C., Brismee, J. M. et Sizer, P. S., Jr. (2006). Subjective and objective descriptors of clinical lumbar spine instability: a Delphi study. *Manual Therapy*, 11(1), 11-21. <https://doi.org/10.1016/j.math.2005.01.002>
- Corbière, M., Bisson, J., Lauzon, S. et Ricard, N. (2006). Factorial validation of a French short-form of the working alliance inventory. *International Journal of Methods in Psychiatric Research*, 15(1), 36-45.
- Costa, L. C., Koes, B. W., Pransky, G., Borkan, J., Maher, C. G. et Smeets, R. J. (2013). Primary care research priorities in low back pain: An update. *Spine*, 38(2), 148-156. <https://doi.org/10.1097/BRS.0b013e318267a92f>
- Coutu, M.-F., Baril, R., Durand, M.-J., Côté, D. et Cadieux, G. (2013). Clinician-patient agreement about the work disability problem of patients having persistent pain: Why it matters. *Journal of Occupational Rehabilitation*, 23(1), 82-92. <https://doi.org/10.1007/s10926-012-9387-8>
- Coutu, M.-F., Dupuis, G. H., Marchand, A., O'Connor, K., Trudel, G. et Bouthillier, D. (2000). Adoption et maintien des habitudes comportementales saines : recension des modèles explicatifs. *Revue francophone de clinique comportementale et cognitive*, 5(2), 23-35.
- Damsgard, E., Thrane, G., Anke, A., Fors, T. et Roe, C. (2010). Activity-related pain in patients with chronic musculoskeletal disorders. *Disability and Rehabilitation*, 32(17), 1428-1437. <https://doi.org/10.3109/09638280903567877>
- Danneels, L. A., Vanderstraeten, G. G., Cambier, D. C., Witvrouw, E. E. et De Cuyper, H. J. (2000). CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. *European Spine Journal*, 9(4), 266-272. <https://doi.org/10.1007/s005860000190>
- Davarian, S., Maroufi, N., Ebrahimi, I., Farahmand, F. et Parnianpour, M. (2012). Trunk muscles strength and endurance in chronic low back pain patients with and without clinical instability. *Journal of Back and Musculoskeletal Rehabilitation*, 25(2), 123-129. <https://doi.org/10.3233/BMR-2012-0320>

- Delitto, A., Erhard, R. E. et Bowling, R. W. (1995). A treatment-based classification approach to low back syndrome: Identifying and staging patients for conservative treatment. *Physical Therapy*, 75(6), 470-485. <https://doi.org/10.1093/ptj/75.6.470>
- Demidenko, E. (2007). Sample size determination for logistic regression revisited. *Statistics in Medicine*, 26(18), 3385-3397. <https://doi.org/10.1002/sim.2771>
- Devilly, G. J. et Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73-86. [https://doi.org/10.1016/S0005-7916\(00\)00012-4](https://doi.org/10.1016/S0005-7916(00)00012-4)
- Deyo, R. A., Rainville, J. et Kent, D. L. (1992). What can the history and physical examination tell us about low back pain? *JAMA*, 268(6), 760-765.
- Dombrowski, M. E., Ryneerson, B., LeVasseur, C., Adgate, Z., Donaldson, W. F., Lee, J. Y., . . . Anderst, W. J. (2018). ISSLS prize in bioengineering science 2018: Dynamic imaging of degenerative spondylolisthesis reveals mid-range dynamic lumbar instability not evident on static clinical radiographs. *European Spine Journal*, 27(4), 752-762. <https://doi.org/10.1007/s00586-018-5489-0>
- Dworkin, R. H., Turk, D. C., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Katz, N. P., . . . Witter, J. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*, 113(1), 9-19. <https://doi.org/10.1016/j.pain.2004.09.012>
- Eisenstein, S. M. (1999). "Instability" and low back pain: A way out of the semantic maze. Dans M. Szpalski, R. Gunzburg et M. H. Pope (édit.), *Lumbar segmental instability* (p. 39-51). Lippincott Williams & Wilkins.
- Enthoven, P., Skargren, E., Kjellman, G. et Oberg, B. (2003). Course of back pain in primary care: A prospective study of physical measures. *Journal of Rehabilitation Medicine*, 35(4), 168-173. <https://doi.org/10.1080/16501970306124>
- Fairbank, J. C. T., Couper, J., Davies, J. et O'Brien, J. (1980). The Oswestry low back pain disability questionnaire. *Physiotherapy*, 66(8), 271-273.
- Faul, F., Erdfelder, E., Buchner, A. et Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Ferreira, P. H., Ferreira, M. L. et Hodges, P. W. (2004). Changes in recruitment of the abdominal muscles in people with low back pain: Ultrasound measurement of muscle activity. *Spine*, 29(22), 2560-2566. <https://doi.org/10.1097/01.brs.0000144410.89182.f9>
- Ferreira, P. H., Ferreira, M. L., Maher, C. G., Herbert, R. D. et Refshauge, K. (2006). Specific stabilisation exercise for spinal and pelvic pain: A systematic review. *Australian Journal of Physiotherapy*, 52(2), 79-88. [https://doi.org/10.1016/s0004-9514\(06\)70043-5](https://doi.org/10.1016/s0004-9514(06)70043-5)

- Ferreira, P. H., Ferreira, M. L., Maher, C. G., Refshauge, K., Herbert, R. D. et Hodges, P. W. (2010). Changes in recruitment of transversus abdominis correlate with disability in people with chronic low back pain. *British Journal of Sports Medicine*, 44(16), 1166-1172. <https://doi.org/10.1136/bjsm.2009.061515>
- Ferrari, S., Manni, T., Bonetti, F., Villafane, J. H. et Vanti, C. (2015). A literature review of clinical tests for lumbar instability in low back pain: Validity and applicability in clinical practice. *Chiropractic & Manual Therapies*, 23, article 14. <https://doi.org/10.1186/s12998-015-0058-7>
- Fitzgerald, G. K. (2010). Invited commentary. *Physical Therapy*, 90(6), 855-858. <https://doi.org/10.2522/ptj.20090233.ic>
- Flor, H., Braun, C., Elbert, T. et Birbaumer, N. (1997). Extensive reorganization of primary somatosensory cortex in chronic back pain patients. *Neuroscience Letters*, 224(1), 5-8. [https://doi.org/10.1016/s0304-3940\(97\)13441-3](https://doi.org/10.1016/s0304-3940(97)13441-3)
- Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., . . . Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, 27(24), 2835-2843.
- Ford, J. J., Bower, S. E., Ford, I., de Mello, M. M., Carneiro, S. R., Balasundaram, A. P. et Hahne, A. J. (2020). Effects of specific muscle activation for low back pain on activity limitation, pain, work participation, or recurrence: A systematic review. *Musculoskeletal Science and Practice*, 50, article 102276. <https://doi.org/10.1016/j.msksp.2020.102276>
- Foster, N. E., Anema, J. R., Cherkin, D., Chou, R., Cohen, S. P., Gross, D. P., . . . Maher, C. G. (2018). Prevention and treatment of low back pain: Evidence, challenges, and promising directions. *The Lancet*, 391(10137), 2368-2383. [https://doi.org/10.1016/S0140-6736\(18\)30489-6](https://doi.org/10.1016/S0140-6736(18)30489-6)
- Foster, N. E., Bishop, A., Thomas, E., Main, C., Horne, R., Weinman, J. et Hay, E. (2008). Illness perceptions of low back pain patients in primary care: What are they, do they change and are they associated with outcome? *Pain*, 136(1-2), 177-187. <https://doi.org/10.1016/j.pain.2007.12.007>
- Freburger, J. K., Carey, T. S., Holmes, G. M., Wallace, A. S., Castel, L. D., Darter, J. D. et Jackman, A. M. (2009). Exercise prescription for chronic back or neck pain: Who prescribes it? Who gets it? What is prescribed? *Arthritis and Rheumatism*, 61(2), 192-200. <https://doi.org/10.1002/art.24234>
- Freburger, J. K., Holmes, G. M., Agans, R. P., Jackman, A. M., Darter, J. D., Wallace, A. S., . . . Carey, T. S. (2009). The rising prevalence of chronic low back pain. *Archives of Internal Medicine*, 169(3), 251-258. <https://doi.org/10.1001/archinternmed.2008.543>
- Freidin, M. B., Tsepilov, Y. A., Stanaway, I. B., Meng, W., Hayward, C., Smith, B. H., . . . Pain, H. A.-I. (2021). Sex- and age-specific genetic analysis of chronic back pain. *Pain*, 162(4), 1176-1187. <https://doi.org/10.1097/j.pain.0000000000002100>

- Fritz, J. M. (2009). Clinical prediction rules in physical therapy: Coming of age? *Journal of Orthopaedic & Sports Physical Therapy*, 39(3), 159-161.
<https://www.jospt.org/doi/10.2519/jospt.2009.0110>
- Fritz, J. M., Delitto, A. et 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(13), 1363-1371.
<https://doi.org/10.1097/01.BRS.0000067115.61673.FF>
- Fritz, J. M., Hebert, J., Koppenhaver, S. et Parent, E. (2009). Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain. *Spine*, 34(25), 2803-2809. <https://doi.org/10.1097/BRS.0b013e3181ae2bd4>
- Fritz, J. M., Piva, S. R. et Childs, J. D. (2005). Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *European Spine Journal*, 14(8), 743-750.
<https://doi.org/10.1007/s00586-004-0803-4>
- Ghezelbash, F., Shahvarpour, A., Larivière, C. et Shirazi-Adl, A. (2021). Evaluating stability of human spine in static tasks: A combined in vivo-computational study. *Journal of Biomechanical Engineering*, (sous presse).
<https://doi.org/10.1080/10255842.2021.2004399>
- Goldby, L. J., Moore, A. P., Doust, J. et Trew, M. E. (2006). A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. *Spine*, 31(10), 1083-1093.
<https://doi.org/10.1097/01.brs.0000216464.37504.64>
- Goldstein, M. S., Elliott, S. D. et Guccione, A. A. (2000). The development of an instrument to measure satisfaction with physical therapy. *Physical Therapy*, 80(9), 853-863.
<https://doi.org/10.1093/ptj/80.9.853>
- Gopika, S. et Dinakaran, R. (2018). Effect of eight-week core stabilization program on core muscle endurance and function in female staff nurses. *National Journal of Physiology, Pharmacy and Pharmacology*, 8(5), 623-626.
<https://doi.org/10.5455/njppp.2018.8.1040407122017>
- Granata, K. P. et England, S. A. (2006). Stability of dynamic trunk movement. *Spine*, 31(10), E271-E276. <http://www.ncbi.nlm.nih.gov/pubmed/16648732>
- Granata, K. P. et Gottipati, P. (2008). Fatigue influences the dynamic stability of the torso. *Ergonomics*, 51(8), 1258-1271. <https://doi.org/10.1080/00140130802030722>
- Granata, K. P. et Orishimo, K. F. (2001). Response of trunk muscle coactivation to changes in spinal stability. *Journal of Biomechanics*, 34(9), 1117-1123.
- Hall, A. M., Ferreira, P. H., Maher, C. G., Latimer, J. et Ferreira, M. L. (2010). The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: A systematic review. *Physical Therapy*, 90(8), 1099-1110.

- Hashemi, L., Webster, B. S., Edward, A. et Enerst, V. (1997). Length of disability and cost of workers' compensation low back pain claims. *Journal of Occupational and Environment Medicine*, 39(10), 937-945.
- Haskins, R., Rivett, D. A. et Osmotherly, P. G. (2012). Clinical prediction rules in the physiotherapy management of low back pain: A systematic review. *Manual Therapy*, 17(1), 9-21. <https://doi.org/10.1016/j.math.2011.05.001>
- Hayden, J. A., van Tulder, M. W. et Tomlinson, G. (2005). Systematic review: Strategies for using exercise therapy to improve outcomes in chronic low back pain. *Annals of Internal Medicine*, 142(9), 776-785. <https://doi.org/10.7326/0003-4819-142-9-200505030-00014>
- Haynes, R. B., Taylor, D. W. et Sackett, D. L. (1979). *Compliance in health care*. Johns Hopkins University Press.
- Hebert, J. J., Koppenhaver, S. L., Magel, J. S. et 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. *Archives of Physical Medicine and Rehabilitation*, 91(1), 78-85. <https://doi.org/10.1016/j.apmr.2009.08.146>
- Henry, S. M. et 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(10), 627-634. <https://doi.org/10.2519/jospt.2007.2555>
- Henschke, N., Maher, C. G., Refshauge, K. M., Das, A. et McAuley, J. H. (2007). Low back pain research priorities: A survey of primary care practitioners. *BMC Family Practice*, 8(1), article 40. <https://doi.org/10.1186/1471-2296-8-40>
- Hicks, G. E. (2011). Invited commentary on "low interrater reliability of examiners performing the prone instability test: A clinical test for lumbar shear instability". *Archives of Physical Medicine and Rehabilitation*, 92(6), 920-922.
- Hicks, G. E., Benvenuti, F., Fiaschi, V., Lombardi, B., Segenni, L., Stuart, M., . . . Macchi, C. (2012). Adherence to a community-based exercise program is a strong predictor of improved back pain status in older adults: An observational study. *Clinical Journal of Pain*, 28(3), 195-203. <https://doi.org/10.1097/AJP.0b013e318226c411>
- Hicks, G. E., Fritz, J. M., Delitto, A. et 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. *Archives of Physical Medicine and Rehabilitation*, 86(9), 1753-1762.
- Hicks, G. E., Fritz, J. M., Delitto, A. et Mishock, J. (2003). Interrater reliability of clinical examination measures for identification of lumbar segmental instability. *Archives of Physical Medicine and Rehabilitation*, 84(12), 1858-1864.

- Hides, J. A., Jull, G. A. et Richardson, C. A. (2001). Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine*, 26(11), E243-248. <https://doi.org/10.1097/00007632-200106010-00004>
- Hides, J. A., Richardson, C. A. et Jull, G. A. (1996). Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. *Spine*, 21(23), 2763-2769. <https://doi.org/10.1097/00007632-199612010-00011>
- Hill, J. C., Dunn, K. M., Lewis, M., Mullis, R., Main, C. J., Foster, N. E. et Hay, E. M. (2008). A primary care back pain screening tool: Identifying patient subgroups for initial treatment. *Arthritis and Rheumatism*, 59(5), 632-641. <https://doi.org/10.1002/art.23563>
- Hodges, P. W. et 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(9), 1005-1012. [https://doi.org/10.1016/s0003-9993\(99\)90052-7](https://doi.org/10.1016/s0003-9993(99)90052-7)
- Hoppes, C. W., Sperier, A. D., Hopkins, C. F., Griffiths, B. D., Principe, M. F., Schnall, B. L., . . . Koppenhaver, S. L. (2016). The efficacy of an eight-week core stabilization program on core muscle function and endurance: A randomized trial. *International Journal of Sports Physical Therapy*, 11(4), 507-519. <https://www.ncbi.nlm.nih.gov/pubmed/27525175>
- Horvath, A. O. et Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36(2), 223-233.
- Hoy, D., March, L., Brooks, P., Blyth, F., Woolf, A., Bain, C., . . . Buchbinder, R. (2014). The global burden of low back pain: Estimates from the Global Burden of Disease 2010 study. *Annals of Rheumatic Diseases*, 73(6), 968-974.
- Hsieh, F. Y. (1989). Sample size tables for logistic regression. *Statistics in Medicine*, 8(7), 795-802. <https://doi.org/10.1002/sim.4780080704>
- Ilfeld, F. W. (1976). Further validation of a psychiatric index in a normal population. *Psychological Reports*, 39, 1215-1228.
- Ito, T., Shirado, O., Suzuki, H., Takahashi, M., Kaneda, K. et Strax, T. E. (1996). Lumbar trunk muscle endurance testing: An inexpensive alternative to a machine for evaluation. *Archives of Physical Medicine and Rehabilitation*, 77, 75-79.
- Iversen, M. D., Fossel, A. H. et Katz, J. N. (2003). Enhancing function in older adults with chronic low back pain: A pilot study of endurance training. *Archives of Physical Medicine and Rehabilitation*, 84(9), 1324-1331. [https://doi.org/10.1016/S0003-9993\(03\)00198-9](https://doi.org/10.1016/S0003-9993(03)00198-9)
- Jack, K., McLean, S. M., Moffett, J. K. et Gardiner, E. (2010). Barriers to treatment adherence in physiotherapy outpatient clinics: A systematic review. *Manual Therapy*, 15(3), 220-228.
- Jaeschke, R., Guyatt, G. H. et Sackett, D. L. (1994). Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients?: The Evidence-Based Medicine Working Group. *JAMA*, 271(9), 703-707. <https://doi.org/10.1001/jama.1994.03510330081039>

- Javadian, Y., Behtash, H., Akbari, M., Taghipour-Darzi, M. et Zekavat, H. (2012). The effects of stabilizing exercises on pain and disability of patients with lumbar segmental instability. *Journal of Back and Musculoskeletal Rehabilitation*, 25(3), 149-155.
<https://doi.org/10.3233/BMR-2012-0321>
- Jensen, M. P., Turner, L. R., Turner, J. A. et Romano, J. M. (1996). The use of multiple-item scales for pain intensity measurement in chronic pain patients. *Pain*, 67(1), 35-40.
- Jones, C. J., Smith, H. E. et Llewellyn, C. D. (2016). A systematic review of the effectiveness of interventions using the Common Sense Self-Regulatory Model to improve adherence behaviours. *Journal of Health Psychology*, 21(11), 2709-2724.
<https://doi.org/10.1177/1359105315583372>
- Jordan, J. L., Holden, M. A., Mason, E. E. J. et Foster, N. E. (2010). Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database Systematic Review*. <https://doi.org/10.1002/14651858.CD005956.pub2>
- Kasai, Y., Morishita, K., Kawakita, E., Kondo, T. et Uchida, A. (2006). A new evaluation method for lumbar spinal instability: Passive lumbar extension test. *Physical Therapy*, 86(12), 1661-1667.
- Kiefer, A., Parnianpour, M. et Shirazi-Adl, A. (1997). Stability of the human spine in neutral postures. *European Spine Journal*, 6(1), 45-53.
- Kobayashi, S., Shizu, N., Suzuki, Y., Asai, T. et Yoshizawa, H. (2003). Changes in nerve root motion and intradiscal blood flow during an intraoperative straight-leg-raising test. *Spine*, 28(13), 1427-1434.
- Koes, B. W., van Tulder, M. W. et Thomas, S. (2006). Diagnosis and treatment of low back pain. *BMJ*, 332(7555), 1430-1434.
- Koumantakis, G. A., Watson, P. J. et Oldham, J. A. (2005). Trunk muscle stabilization training plus general exercise versus general exercise only: Randomized controlled trial of patients with recurrent low back pain. *Physical Therapy*, 85(3), 209-225.
- L'Heureux, J., Coutu, M.-F., Berbiche, D. et Larivière, C. (2019). Adherence to a home exercise programme following a clinical programme for non-acute non-specific low back pain: An exploratory study. *European Journal of Physiotherapy*, 22(5), 299-308.
<https://doi.org/10.1080/21679169.2019.1617777>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, article 863.
<https://doi.org/10.3389/fpsyg.2013.00863>
- Lamoth, C. J., Daffertshofer, A., Meijer, O. G. et Beek, P. J. (2006). How do persons with chronic low back pain speed up and slow down?: Trunk-pelvis coordination and lumbar erector spinae activity during gait. *Gait & Posture*, 23(2), 230-239.
<https://doi.org/10.1016/j.gaitpost.2005.02.006>

- Larivière, C., Boucher, J. A., Mecheri, H. et Ludvig, D. (2019). Maintaining lumbar spine stability: A study of the specific and combined effects of abdominal activation and lumbosacral orthosis on lumbar intrinsic stiffness. *Journal of Orthopaedic & Sports Physical Therapy*, 49(4), 262-271. <https://www.jospt.org/doi/10.2519/jospt.2019.8565>
- Larivière, C., Butler, H., Sullivan, M. J. et Fung, J. (2013). An exploratory study on the effect of pain interference and attentional interference on neuromuscular responses during rapid arm flexion movements. *Clinical Journal of Pain*, 29(3), 265-275.
- Larivière, C., Coutu, M. F., Henry, S. M., Preuss, R., Gagnon, D., Dumas, J. P. et Sullivan, M. J. L. (2016). *Développement préliminaire d'une règle de prédiction clinique pour dépister les patients ayant une lombalgie non-aigüe répondant favorablement à un programme d'exercices de stabilisation lombaire* (Rapport n° R-935). IRSST.
- Larivière, C., Gagnon, D. H., Henry, S. M., Preuss, R. et Dumas, J. P. (2018). The effects of an 8-week stabilization exercise program on lumbar multifidus muscle thickness and activation as measured with ultrasound imaging in patients with low back pain: An exploratory study. *PM&R*, 10(5), 483-493. <https://doi.org/10.1016/j.pmrj.2017.10.005>
- Larivière, C., Henry, S. M., Gagnon, D. H., Preuss, R. et Dumas, J. P. (2019). Ultrasound measures of the abdominal wall in patients with low back pain before and after an 8-week lumbar stabilization exercise program, and their association with clinical outcomes. *PM&R*, 11(7), 710-721. <https://doi.org/10.1002/pmrj.12000>
- Larivière, C., Henry, S. M. et Preuss, R. (2020). Structural remodeling of the lumbar multifidus, thoracolumbar fascia and lateral abdominal wall perimuscular connective tissues: A search for its potential determinants. *Journal of Anatomy*, 238(3), 536-550. <https://doi.org/10.1111/joa.13330>
- Larivière, C., Ludvig, D., Kearney, R. E., Mecheri, H., Caron, J.-M. et Preuss, R. (2015). Identification of intrinsic and reflexive contributions to low-back stiffness: Medium-term reliability and construct validity. *Journal of Biomechanics*, 48, 254-261.
- Larivière, C., Mecheri, H., Shahvarpour, A., Gagnon, D. et Shirazi-Adl, A. (2013). Criterion validity and between-day reliability of an inertial-sensor-based trunk postural stability test during unstable sitting. *Journal of Electromyography and Kinesiology*, 23(4), 899-907. <https://doi.org/10.1016/j.jelekin.2013.03.002>
- Larivière, C., Preuss, R., Gagnon, D. H., Mecheri, H. et Henry, S. M. (2020). Structural remodelling of the lumbar multifidus, thoracolumbar fascia and lateral abdominal wall perimuscular connective tissues: A cross-sectional and comparative ultrasound study. *Journal of Bodywork and Movement Therapies*, 24(4), 293-302. <https://doi.org/10.1016/j.jbmt.2020.07.009>
- Larivière, C., Shahvarpour, A., Gravel, C., Gauvin, M., Jean, A.-M., Viau, A. et Mecheri, H. (2019). Revisiting the effect of manipulating lumbar stability with load magnitudes and positions: The effect of sex on trunk muscle activation. *Journal of Electromyography and Kinesiology*, 46, 55-62.

- Lebeau, M. (2021). *Valeurs estimées par le Groupe connaissance et surveillance statistique (GCSS) de la Direction de la recherche de l'IRSST (CNESS; 2016-2018)*. Données non publiées. IRSST.
- Leboeuf-Yde, C. et Manniche, C. (2001). Low back pain: Time to get off the treadmill. *Journal of Manipulative and Physiological Therapeutics*, 24(1), 63-66.
<https://doi.org/10.1067/mmt.2001.112009>
- Lee, A. S., Cholewicki, J., Reeves, N. P., Zazulak, B. T. et Mysliwiec, L. W. (2010). Comparison of trunk proprioception between patients with low back pain and healthy controls. *Archives of Physical Medicine and Rehabilitation*, 91(9), 1327-1331.
- Lee, H. S., Kim, D. J., Oh, Y., Min, K. et Ryu, J. S. (2016). The effect of individualized gradable stabilization exercises in patients with chronic low back pain: Case-control study. *Journal of Back and Musculoskeletal Rehabilitation*, 29(3), 603-610.
<https://doi.org/10.3233/BMR-160724>
- Lee, W. C. (1999). Selecting diagnostic tests for ruling out or ruling in disease: The use of the Kullback-Leibler distance. *International Journal of Epidemiology*, 28(3), 521-525.
<https://doi.org/10.1093/ije/28.3.521>
- Leventhal, H. (1980). The role of theory in the study of compliance to high blood pressure regimens. Dans B. R. M. Haynes, M. E. et Engebretson, T. O. (édit.), *Patient compliance to prescribed antihypertensive medication regimens: A report to the National Heart, Lung, and Blood Institute*. National Heart, Lung, and Blood Institute.
- Leventhal, H., Owa, R. D. et Brissette, I. (2012). The common-sense model of self-regulation of health and illness. Dans L. Cameron et H. Leventhal (édit.), *The self-regulation of health and illness behaviour* (p. 56-79). Routledge.
- Liddle, S. D., Baxter, G. D. et Gracey, J. H. (2004). Exercise and chronic low back pain: What works? *Pain*, 107(1-2), 176-190.
- Liebenson, C., Karpowicz, A. M., Brown, S. H., Howarth, S. J. et McGill, S. M. (2009). The active straight leg raise test and lumbar spine stability. *PM&R*, 1(6), 530-535.
<https://doi.org/10.1016/j.pmrj.2009.03.007>
- Long, A., Donelson, R. et Fung, T. (2004). Does it matter which exercise?: A randomized control trial of exercise for low back pain. *Spine*, 29(23), 2593-2602.
<https://doi.org/10.1097/01.brs.0000146464.23007.2a>
- Luomajoki, H., Kool, J., de Bruin, E. D. et Airaksinen, O. (2007). Reliability of movement control tests in the lumbar spine. *BMC Musculoskeletal Disorders*, 8, article 90.
<https://doi.org/10.1186/1471-2474-8-90>
- Lyles, R. H., Lin, H. M. et Williamson, J. M. (2007). A practical approach to computing power for generalized linear models with nominal, count, or ordinal responses. *Statistics in Medicine*, 26(7), 1632-1648. <https://doi.org/10.1002/sim.2617>

- MacDonald, D., Moseley, G. L. et 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(3), 183-188. <https://doi.org/10.1016/j.pain.2008.12.002>
- MacDonald, D., Moseley, G. L. et Hodges, P. W. (2010). People with recurrent low back pain respond differently to trunk loading despite remission from symptoms. *Spine*, 35(7), 818-824. <https://doi.org/10.1097/BRS.0b013e3181bc98f1>
- Macedo, L. G., Maher, C. G., Latimer, J. et McAuley, J. H. (2009). Motor control exercise for persistent, nonspecific low back pain: A systematic review. *Physical Therapy*, 89(1), 9-25. <https://doi.org/10.2522/ptj.20080103>
- Mannion, A. F., Helbling, D., Pulkovski, N. et Sprott, H. (2009). Spinal segmental stabilisation exercises for chronic low back pain: Programme adherence and its influence on clinical outcome. *European Spine Journal*, 18(12), 1881-1891. <https://doi.org/10.1007/s00586-009-1093-7>
- Mannion, A. F., Kaser, L., Weber, E., Rhyner, A., Dvorak, J. et Muntener, M. (2000). Influence of age and duration of symptoms on fibre type distribution and size of the back muscles in chronic low back pain patients. *European Spine Journal*, 9(4), 273-281. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3611339/pdf/586_2000_Article_90273.586.pdf
- Mannion, A. F., Muntener, M., Taimela, S. et Dvorak, J. (1999). A randomized clinical trial of three active therapies for chronic low back pain. *Spine*, 24(23), 2435-2448.
- Marcus, B. H., Selby, V. C., Niaura, R. S. et Rossi, J. S. (1992). Self-efficacy and the stages of exercise behavior change. *Research Quarterly for Exercise and Sport*, 63(1), 60-66. <https://doi.org/10.1080/02701367.1992.10607557>
- Marshall, P. et Murphy, B. (2006). Changes in the flexion relaxation response following an exercise intervention. *Spine*, 31(23), E877-883. <https://doi.org/10.1097/01.brs.0000244557.56735.05>
- Marshall, P. W., Kennedy, S., Brooks, C. et Lonsdale, C. (2013). Pilates exercise or stationary cycling for chronic nonspecific low back pain: Does it matter?: A randomized controlled trial with 6-month follow-up. *Spine*, 38(15), E952-E959. <https://doi.org/10.1097/BRS.0b013e318297c1e5>
- Masse-Alarie, H., Beaulieu, L. D., Preuss, R. et Schneider, C. (2016a). Corticomotor control of lumbar multifidus muscles is impaired in chronic low back pain: Concurrent evidence from ultrasound imaging and double-pulse transcranial magnetic stimulation. *Experimental Brain Research*, 234(4), 1033-1045. <https://doi.org/10.1007/s00221-015-4528-x>
- Masse-Alarie, H., Beaulieu, L. D., Preuss, R. et Schneider, C. (2016b). Influence of chronic low back pain and fear of movement on the activation of the transversely oriented abdominal muscles during forward bending. *Journal of Electromyography and Kinesiology*, 27, 87-94. <https://doi.org/10.1016/j.jelekin.2016.02.004>

- Masse-Alarie, H., Beaulieu, L. D., Preuss, R. et Schneider, C. (2017). The side of chronic low back pain matters: Evidence from the primary motor cortex excitability and the postural adjustments of multifidi muscles. *Experimental Brain Research*, 235(3), 647-659. <https://doi.org/10.1007/s00221-016-4834-y>
- McAuley, E. (1992). *Barriers Self-Efficacy Scale (BARSE)*. <http://www.epl.illinois.edu/measures/barse.html>
- McAuley, E., Jerome, G. J., Marquez, D. X., Elavsky, S. et Blissmer, B. (2003). Exercise self-efficacy in older adults: Social, affective, and behavioral influences. *Annals of Behavioral Medicine*, 25(1), 1-7. https://doi.org/10.1207/S15324796ABM2501_01
- McCarthy, C. J., Arnall, F. A., Strimpakos, N., Freemont, A. et Oldham, J. A. (2004). The biopsychosocial classification of non-specific low back pain: A systematic review. *Physical Therapy Reviews*, 9(1), 17-30. <https://doi.org/10.1179/108331904225003955>
- McCarthy, C. J., Gittins, M., Roberts, C. et Oldham, J. A. (2007). The reliability of the clinical tests and questions recommended in international guidelines for low back pain. *Spine*, 32(8), 921-926. <https://doi.org/10.1097/01.brs.0000259864.21869.26>
- McCracken, L. M., Gross, R. T. et Eccleston, C. (2002). Multimethod assessment of treatment process in chronic low back pain: Comparison of reported pain-related anxiety with directly measured physical capacity. *Behaviour Research Therapy*, 40(5), 585-594. [https://doi.org/10.1016/s0005-7967\(01\)00074-2](https://doi.org/10.1016/s0005-7967(01)00074-2)
- McGill, S. (2002). *Low back disorders: Evidence-based prevention and rehabilitation*. Human Kinetics.
- McGill, S. M., Childs, A. et Liebenson, C. (1999). Endurance times of low back stabilization exercises: Clinical targets for testing and training from a normal database. *Archives of Physical Medicine and Rehabilitation*, 80(8), 941-944.
- McLean, S. M., Burton, M., Bradley, L. et Littlewood, C. (2010). Interventions for enhancing adherence with physiotherapy: A systematic review. *Manual Therapy*, 15(6), 514-521.
- Mehta, R., Cannella, M., Henry, S. M., Smith, S., Giszter, S. et Silfies, S. P. (2017). Trunk postural muscle timing is not compromised in low back pain patients clinically diagnosed with movement coordination impairments. *Motor Control*, 21(2), 133-157. <https://doi.org/10.1123/mc.2015-0049>
- Mens, J. M., Vleeming, A., Snijders, C. J., Koes, B. W. et Stam, H. J. (2001). Reliability and validity of the active straight leg raise test in posterior pelvic pain since pregnancy. *Spine*, 26(10), 1167-1171.
- Moseley, G. L. et Hodges, P. W. (2006). Reduced variability of postural strategy prevents normalization of motor changes induced by back pain: A risk factor for chronic trouble? *Behavioral Neuroscience*, 120(2), 474-476. <https://doi.org/10.1037/0735-7044.120.2.474>

- Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D. et Buick, D. (2002). The revised illness perception questionnaire (IPQ-R). *Psychology and Health*, 17(1), 1-16.
- Nava-Bringas, T. I., Roeniger-Desatnik, A., Arellano-Hernández, A. et Cruz-Medina, E. (2016). Adherence to a stability exercise programme in patients with chronic low back pain. *Cirugía y Cirujanos (English Edition)*, 84(5), 384-391.
- Ng, J. K., Kippers, V., Richardson, C. A. et Parnianpour, M. (2001). Range of motion and lordosis of the lumbar spine: Reliability of measurement and normative values. *Spine*, 26(1), 53-60.
- Niederer, D. et Mueller, J. (2020). Sustainability effects of motor control stabilisation exercises on pain and function in chronic nonspecific low back pain patients: A systematic review with meta-analysis and meta-regression. *PLoS One*, 15(1), article e0227423. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6961919/pdf/pone.0227423.pdf>
- Novy, D. M., Simmonds, M. J. et Lee, C. E. (2002). Physical performance tasks: What are the underlying constructs? *Archives of Physical Medicine and Rehabilitation*, 83(1), 44-47. <https://doi.org/10.1053/apmr.2002.27397>
- O'Sullivan, P. (2005). Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. *Manual Therapy*, 10(4), 242-255.
- O'Sullivan, P. B., Phyt, G. D., Twomey, L. T. et 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(24), 2959-2967. <https://doi.org/10.1097/00007632-199712150-00020>
- Ostelo, R. W., Deyo, R. A., Stratford, P., Waddell, G., Croft, P., Von Korf, M., . . . de Vet, H. C. (2008). Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine*, 33(1), 90-94.
- Panjabi, M. M. (1992a). The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. *Journal of Spinal Disorders*, 5(4), 383-389.
- Panjabi, M. M. (1992b). The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. *Journal of Spinal Disorders*, 5(4), 390-397.
- Panjabi, M. M. (2006). A hypothesis of chronic back pain: Ligament subfailure injuries lead to muscle control dysfunction. *European Spine Journal*, 15(5), 668-676.
- Park, S. H. et Lee, M. M. (2021). Effects of progressive neuromuscular stabilization exercise on the support surface on patients with high obesity with lumbar instability: A double-blinded randomized controlled trial. *Medicine*, 100(4), article e23285. <https://doi.org/10.1097/MD.00000000000023285>

- Peek, K., Sanson-Fisher, R., Mackenzie, L. et Carey, M. (2016). Interventions to aid patient adherence to physiotherapist prescribed self-management strategies: A systematic review. *Physiotherapy*, 102(2), 127-135. <https://doi.org/10.1016/j.physio.2015.10.003>
- Préville, M., Boyer, R. et Potvin, L. (1992). *La détresse psychologique : détermination de la fiabilité et de la validité de la mesure utilisée dans l'Enquête Santé Québec*. Santé Québec.
- Puntumetakul, R., Yodchaisarn, W., Emasithi, A., Keawduangdee, P., Chatchawan, U. et Yamauchi, J. (2015). Prevalence and individual risk factors associated with clinical lumbar instability in rice farmers with low back pain. *Patient Preference and Adherence*, 9, 1-7.
- Rabin, A., Shashua, A., Pizem, K. et Dar, G. (2013). The interrater reliability of physical examination tests that may predict the outcome or suggest the need for lumbar stabilization exercises. *Journal of Orthopaedic & Sports Physical Therapy*, 43(2), 83-90.
- Rabin, A., Shashua, A., Pizem, K., Dickstein, R. et Dar, G. (2014). A clinical prediction rule to identify patients with low back pain who are likely to experience short-term success following lumbar stabilization exercises: A randomized controlled validation study. *Journal of Orthopaedic & Sports Physical Therapy*, 44(1), 6-B13. <https://doi.org/10.2519/jospt.2014.4888>
- Ravenna, M. M., Hoffman, S. L. et Van Dillen, L. R. (2011). Low interrater reliability of examiners performing the prone instability test: A clinical test for lumbar shear instability. *Archives of Physical Medicine and Rehabilitation*, 92(6), 913-919. <https://doi.org/10.1016/j.apmr.2010.12.042>
- Reeves, N. P., Cholewicki, J., van Dieen, J. H., Kawchuk, G. et Hodges, P. W. (2019). Are stability and instability relevant concepts for back pain? *Journal of Orthopaedic & Sports Physical Therapy*, 49(6), 415-424. <https://doi.org/10.2519/jospt.2019.8144>
- Reeves, N. P., Narendra, K. S. et Cholewicki, J. (2007). Spine stability: The six blind men and the elephant. *Clinical Biomechanics*, 22(3), 266-274. <https://doi.org/10.1016/j.clinbiomech.2006.11.011>
- Reeves, N. P., Narendra, K. S. et Cholewicki, J. (2011). Spine stability: Lessons from balancing a stick. *Clinical Biomechanics*, 26(4), 325-330. <https://doi.org/10.1016/j.clinbiomech.2010.11.010>
- Remvig, L., Jensen, D. V. et Ward, R. C. (2007a). Are diagnostic criteria for general joint hypermobility and benign joint hypermobility syndrome based on reproducible and valid tests?: A review of the literature. *Journal of Rheumatology*, 34(4), 798-803.
- Remvig, L., Jensen, D. V. et Ward, R. C. (2007b). Epidemiology of general joint hypermobility and basis for the proposed criteria for benign joint hypermobility syndrome: Review of the literature. *Journal of Rheumatology*, 34(4), 804-809.

- Resnick, B., Zimmerman, S. I., Orwig, D., Furstenberg, A. L. et Magaziner, J. (2000). Outcome expectations for exercise scale: Utility and psychometrics. *Journals of Gerontology: Series B*, 55(6), S352-S356.
- Richardson, C., Hodges, P. et Hides, J. (2004). *Therapeutic exercise for lumbopelvic stabilization* (2^e éd.). Churchill Livingstone.
- Roland, M., Waddell, G., Klüber Moffett, J., Burton, K. et Main, C. (2009). Le guide du dos (2^e éd.). TSO, trad.).
- Rossignol, M. (édit.). (2006). *CLIP : Clinique des lombalgies interdisciplinaire en première ligne : guide pratique*. Agence de la santé et des services sociaux de Montréal.
- Roussel, N. A., Nijs, J., Truijen, S., Smeuninx, L. et Stassijns, G. (2007). Low back pain: Clinimetric properties of the Trendelenburg test, active straight leg raise test, and breathing pattern during active straight leg raising. *Journal of Manipulative and Physiological Therapeutics*, 30(4), 270-278. <https://doi.org/10.1016/j.jmpt.2007.03.001>
- Sahrmann, S. A. (2002). *Diagnosis and treatment of movement impairment syndromes*. Mosby.
- Sallis, J. F., Grossman, R. M., Pinski, R. B., Patterson, T. L. et Nader, P. R. (1987). The development of scales to measure social support for diet and exercise behaviors. *Preventive Medicine*, 16(6), 825-836. [https://doi.org/10.1016/0091-7435\(87\)90022-3](https://doi.org/10.1016/0091-7435(87)90022-3)
- Sandrey, M. A. et Mitzel, J. G. (2013). Improvement in dynamic balance and core endurance after a 6-week core-stability-training program in high school track and field athletes. *Journal of Sport Rehabilitation*, 22(4), 264-271. <https://doi.org/10.1123/jsr.22.4.264>
- Saragiotto, B. T., Maher, C. G., Yamato, T. P., Costa, L. O., Costa, L. C., Ostelo, R. W. et Macedo, L. G. (2016). Motor control exercise for nonspecific low back pain: A Cochrane Review. *Spine*, 41(16), 1284-1295. <https://doi.org/10.1097/BRS.0000000000001645>
- Saur, P. M. M., Ensink, R. B. M., Frese, K., Seeger, D. et Hildebrandt, J. (1996). Lumbar range of motion: Reliability and validity of the inclinometer technique in the clinical measurement of trunk flexibility. *Spine*, 21(11), 1332-1338.
- Savigny, P., Watson, P. et Underwood, M. (2009). Early management of persistent non-specific low back pain: Summary of NICE guidance. *BMJ*, 338, article b1805. <https://doi.org/10.1136/bmj.b1805>
- Schneider, M., Erhard, R., Brach, J., Tellin, W., Imbarlina, F. et Delitto, A. (2008). Spinal palpation for lumbar segmental mobility and pain provocation: An interexaminer reliability study. *Journal of Manipulative and Physiological Therapeutics*, 31(6), 465-473.
- Searle, A., Spink, M., Ho, A. et Chuter, V. (2015). Exercise interventions for the treatment of chronic low back pain: A systematic review and meta-analysis of randomised controlled trials. *Clinical Rehabilitation*, 29(12), 1155-1167. <https://doi.org/10.1177/0269215515570379>

- Shacklock, M. (1995). Neurodynamics. *Physiotherapy*, 81(1), 9-16.
- Shahvarpour, A., Gagnon, D., Preuss, R., Henry, S. M. et Larivière, C. (2018). Trunk postural balance and low back pain: Reliability and relationship with clinical changes following a lumbar stabilization exercise program. *Gait & Posture*, 61, 375-381.
<https://doi.org/10.1016/j.gaitpost.2018.02.006>
- Shahvarpour, A., Henry, S. M., Preuss, R., Mecheri, H. et Larivière, C. (2017). The effect of an 8-week stabilization exercise program on the lumbopelvic rhythm and flexion-relaxation phenomenon. *Clinical Biomechanics*, 48, 1-8.
<https://doi.org/10.1016/j.clinbiomech.2017.06.010>
- Shahvarpour, A., Shirazi-Adl, A., Larivière, C. et Bazrgari, B. (2015). Computation of trunk stability in forward perturbations: Effects of preload, perturbation load, initial flexion and abdominal preactivation. *Journal of Biomechanics*, 48, 716-720.
- Simmonds, M. J. (2006). Measuring and managing pain and performance. *Manual Therapy*, 11(3), 175-179. <https://doi.org/10.1016/j.math.2006.03.002>
- Simmonds, M. J., Olson, S. L., Jones, S., Hussein, T., Lee, C. E., Novy, D. et Radwan, H. (1998). Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine*, 23(22), 2412-2421.
- Smeets, R. J., Vlaeyen, J. W., Kester, A. D. et Knottnerus, J. A. (2006). Reduction of pain catastrophizing mediates the outcome of both physical and cognitive-behavioral treatment in chronic low back pain. *Journal of Pain*, 7(4), 261-271.
- Smidt, N., de Vet, H. C., Bouter, L. M., Dekker, J., Arendzen, J. H., De Bie, R. A., . . . van der Windt, D. A. (2005). Effectiveness of exercise therapy: A best-evidence summary of systematic reviews. *Australian Journal of Physiotherapy*, 51(2), 71-85.
[https://doi.org/10.1016/S0004-9514\(05\)70036-2](https://doi.org/10.1016/S0004-9514(05)70036-2)
- Smith, B. E., Littlewood, C. et May, S. (2014). An update of stabilisation exercises for low back pain: A systematic review with meta-analysis. *BMC Musculoskeletal Disorders*, 15(1), article 416. <https://doi.org/10.1186/1471-2474-15-416>
- Smith, T. O., Jerman, E., Easton, V., Bacon, H., Armon, K., Poland, F. et Macgregor, A. J. (2013). Do people with benign joint hypermobility syndrome (BJHS) have reduced joint proprioception?: A systematic review and meta-analysis. *Rheumatology International*, 33(11), 2709-2716.
- Stanton, T. R., Fritz, J. M., Hancock, M. J., Latimer, J., Maher, C. G., Wand, B. M. et Parent, E. C. (2011). Evaluation of a treatment-based classification algorithm for low back pain: A cross-sectional study. *Physical Therapy*, 91(4), 496-509.
<https://doi.org/10.2522/ptj.20100272>
- Stanton, T. R., Hancock, M. J., Maher, C. G. et Koes, B. W. (2010). Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. *Physical Therapy*, 90(6), 843-854.

- Steiger, F., Wirth, B., de Bruin, E. D. et Mannion, A. F. (2012). Is a positive clinical outcome after exercise therapy for chronic non-specific low back pain contingent upon a corresponding improvement in the targeted aspect(s) of performance?: A systematic review. *European Spine Journal*, 21(4), 575-598. <https://doi.org/10.1007/s00586-011-2045-6>
- Stolz, M., von Piekartz, H., Hall, T., Schindler, A. et Ballenberger, N. (2019). Evidence and recommendations for the use of segmental motion testing for patients with LBP: A systematic review. *Musculoskeletal Science and Practice*, 45, article 102076. <https://doi.org/10.1016/j.msksp.2019.102076>
- Stolze, L. R., Allison, S. C. et Childs, J. D. (2012). Derivation of a preliminary clinical prediction rule for identifying a sub-group of patients with low back pain likely to benefit from Pilates-based exercise. *Journal of Orthopaedic & Sports Physical Therapy*, 42(5), 425-436. <https://www.jospt.org/doi/10.2519/jospt.2012.3826>
- Storheim, K., Brox, J. I., Holm, I., Koller, A. K. et Bo, K. (2003). Intensive group training versus cognitive intervention in sub-acute low back pain: Short-term results of a single-blind randomized controlled trial. *Journal of Rehabilitation Medicine*, 35(3), 132-140. <https://doi.org/10.1080/16501970310010484>
- Strutton, P. H., Theodorou, S., Catley, M., McGregor, A. H. et Davey, N. J. (2005). Corticospinal excitability in patients with chronic low back pain. *Journal of Spinal Disorders & Techniques*, 18(5), 420-424. <https://doi.org/10.1097/01.bsd.0000169063.84628.fe>
- Sullivan, M. J. L., Bishop, S. R. et Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychological Assessment*, 7, 524-532.
- Sung, W., Hicks, G. E., Ebaugh, D., Smith, S. S., Stackhouse, S., Wattananon, P. et Silfies, S. P. (2019). Individuals with and without low back pain use different motor control strategies to achieve spinal stiffness during the prone instability test. *Journal of Orthopaedic & Sports Physical Therapy*, 49(12), 899-907. <https://doi.org/10.2519/jospt.2019.8577>
- Teyhen, D. S., Flynn, T. W., Childs, J. D. et Abraham, L. D. (2007). Arthrokinematics in a subgroup of patients likely to benefit from a lumbar stabilization exercise program. *Physical Therapy*, 87(3), 313-325. <https://doi.org/10.2522/ptj.20060253>
- Thompson, E., Broadbent, J., Bertino, M. D. et Staiger, P. K. (2015). Do pain-related beliefs influence treatment adherence?: A systematic review. *Clinical Journal of Pain*, 32(2), 164-178. <https://doi.org/10.1097/AJP.0000000000000235>
- Tsao, H., Galea, M. P. et Hodges, P. W. (2008). Reorganization of the motor cortex is associated with postural control deficits in recurrent low back pain. *Brain*, 131(Pt 8), 2161-2171. <https://doi.org/10.1093/brain/awn154>
- Tsao, H., Galea, M. P. et Hodges, P. W. (2009). How fast are feedforward postural adjustments of the abdominal muscles? *Behavioral Neuroscience*, 123(3), 687-693. <https://doi.org/10.1037/a0015593>

- Tsao, H., Galea, M. P. et Hodges, P. W. (2010). Driving plasticity in the motor cortex in recurrent low back pain. *European Journal of Pain*, 14(8), 832-839. <https://doi.org/10.1016/j.ejpain.2010.01.001>
- Unsgaard-Tondel, M., Fladmark, A. M., Salvesen, O. et Vasseljen, O. (2010). Motor control exercises, sling exercises, and general exercises for patients with chronic low back pain: A randomized controlled trial with 1-year follow-up. *Physical Therapy*, 90(10), 1426-1440. <https://doi.org/10.2522/ptj.20090421>
- van Dieen, J. H., Flor, H. et Hodges, P. W. (2017). Low-back pain patients learn to adapt motor behavior with adverse secondary consequences. *Exercise and Sport Sciences Reviews*, 45(4), 223-229. <https://doi.org/10.1249/JES.0000000000000121>
- Van Dillen, L. R., Sahrmann, S. A., Norton, B. J., Caldwell, C. A., Fleming, D., McDonnell, M. K. et Bloom, N. J. (2001). Effect of active limb movements on symptoms in patients with low back pain. *Journal of Orthopaedic & Sports Physical Therapy*, 31(8), 402-413. <https://doi.org/10.2519/jospt.2001.31.8.402>
- Van Dillen, L. R., Sahrmann, S. A., Norton, B. J., Caldwell, C. A., Fleming, D. A., McDonnell, M. K. et Woolsey, N. B. (1998). Reliability of physical examination items used for classification of patients with low back pain. *Physical Therapy*, 78(9), 979-988.
- Van Dillen, L. R., Sahrmann, S. A., Norton, B. J., Caldwell, C. A., McDonnell, M. K. et Bloom, N. J. (2003). Movement system impairment-based categories for low back pain: Stage 1 validation. *Journal of Orthopaedic & Sports Physical Therapy*, 33(3), 126-142. <https://www.jospt.org/doi/10.2519/jospt.2003.33.3.126>
- van Meulenbroek, T., Huijnen, I. P. J., Simons, L. E., Conijn, A. E. A., Engelbert, R. H. H. et Verbunt, J. A. (2020). Exploring the underlying mechanism of pain-related disability in hypermobile adolescents with chronic musculoskeletal pain. *Scandinavian Journal of Pain*, 21(1), 22-31. <https://doi.org/10.1515/sjpain-2020-0023>
- van Tulder, M., Malmivaara, A., Esmail, R. et Koes, B. (2000). Exercise therapy for low back pain: A systematic review within the framework of the Cochrane collaboration back review group. *Spine*, 25(21), 2784-2796.
- Vanti, C., Conti, C., Faresin, F., Ferrari, S. et Piccarreta, R. (2016). The relationship between clinical instability and endurance tests, pain, and disability in nonspecific low back pain. *Journal of Manipulative and Physiological Therapeutics*, 39(5), 359-368. <https://doi.org/10.1016/j.jmpt.2016.04.003>
- Vera-Garcia, F. J., Elvira, J. L., Brown, S. H. et McGill, S. M. (2007). Effects of abdominal stabilization maneuvers on the control of spine motion and stability against sudden trunk perturbations. *Journal of Electromyography and Kinesiology*, 17(5), 556-567. <https://doi.org/10.1016/j.jelekin.2006.07.004>
- Vlaeyen, J. W. S. et Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: A state of the art. *Pain*, 85(3), 317-332. [https://doi.org/10.1016/S0304-3959\(99\)00242-0](https://doi.org/10.1016/S0304-3959(99)00242-0)

- Vuori, I. M. (2001). Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. *Medicine & Science in Sports & Exercise*, 33(6 Suppl), S551-S586.
- Waddell, G., Burton, A. K. et Main, C. J. (2003). *Screening to identify people at risk of long-term incapacity for work: A conceptual and scientific review*. The Royal Society of Medicine Press.
- Waddell, G., Newton, M., Henderson, I., Somerville, D. et Main, C. J. (1993). A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, 52, 157-168.
- Waddell, G., Somerville, D., Henderson, I. et Newton, M. (1992). Objective clinical evaluation of physical impairment in chronic low back pain. *Spine*, 17(6), 617-628. <https://doi.org/10.1097/00007632-199206000-00001>
- Woby, S. R., Watson, P. J., Roach, N. K. et Urmston, M. (2004). Are changes in fear-avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic low back pain and disability? *European Journal of Pain*, 8(3), 201-210. <https://doi.org/10.1016/j.ejpain.2003.08.002>
- Wu, M., Wang, S., Driscoll, S. J., Cha, T. D., Wood, K. B. et Li, G. (2014). Dynamic motion characteristics of the lower lumbar spine: Implication to lumbar pathology and surgical treatment. *European Spine Journal*, 23(11), 2350-2358. <https://doi.org/10.1007/s00586-014-3316-9>
- Wun, A., Kollias, P., Jeong, H., Rizzo, R. R., Cashin, A. G., Bagg, M. K., . . . Jones, M. D. (2020). Why is exercise prescribed for people with chronic low back pain?: A review of the mechanisms of benefit proposed by clinical trialists. *Musculoskeletal Science and Practice*, 51, article 102307. <https://doi.org/10.1016/j.msksp.2020.102307>
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1), 32-35.

ANNEXE A. Description sommaire des mesures

Tableau A-1. Résultats cliniques (OUT) et mesures de l'évaluation psychologique (PSY)

CONCEPT	QUESTIONNAIRE	VARIABLE (abréviation)	USAGE*	
Mesures de résultats cliniques (mesures OUT) – 20 min			Obj.1	Obj.2
Perception d'incapacité	Questionnaire Oswestry (Fairbank <i>et al.</i> , 1980)	OSW	G	G
Intensité de la douleur	Échelle numérique de 11 points (0 : pas de douleur à 10 : insupportable) pour évaluer la douleur moyenne (habituelle) ainsi que la plus petite et la pire douleur durant la dernière semaine (Cleeland et Ryan, 1994). La moyenne des trois scores permet d'en augmenter la validité (Jensen <i>et al.</i> , 1996). L'intensité actuelle (présentement) n'a pas été mesurée, car peut fluctuer en fonction d'une activité effectuée juste avant l'évaluation.	Int-Douleur	P-C	G
Changement global	Échelle de 7 points (-3 [beaucoup aggravé] à +3 [-beaucoup amélioré]) d'estimation globale du changement (Dworkin <i>et al.</i> , 2005).	ChangGlob	R	/
Satisfaction du traitement	Questionnaire de satisfaction du patient sur les services de physiothérapie (Goldstein <i>et al.</i> , 2000)	Satisfaction-PT	R	/
Évaluation psychologique (mesures PSY) – 60 min			Obj.1	Obj.2
Peurs et croyances	<i>Fear-avoidance beliefs questionnaire</i> - 16 items (5 min) (Waddell <i>et al.</i> , 1993). Score/24 (composante activité physique de 5 items)	Peurs/croyances-AP	P-B	R
Dramatisation de la douleur	<i>Pain Catastrophizing Scale</i> – 13 items (5 min) (Sullivan <i>et al.</i> , 1995) Score/52	Dramatisation/douleur	P-B	R
Détresse psychologique	Inventaire de détresse psychologique (IDP) – 29 items (10 min) (Ilfeld, 1976; Prévile <i>et al.</i> , 1992). Score/100. Le score total (IDPtot) n'a pas été considéré comme candidat à la RPC, car il implique 29 items, ce qui a été jugé trop long à l'examen clinique. Les sous-échelles comportent moins d'items et mesurent des concepts plus précis et utiles pour l'intervention.	IDPtot (score total) IDPcog (score cognition) IDPanx (score anxiété) IDPdep (score dépression) IDPcol (score colère)	P-B	R
Douleur liée à l'activité	Est-ce que vous avez une augmentation de la douleur durant vos activités en général ou lors de l'exercice ? O/N (Damsgard <i>et al.</i> , 2010)	DoulActPhys	P-B	R
Niveau d'activité physique (NAP)	Questionnaire sur l'activité physique (dernier mois) – 10 min (Baecke <i>et al.</i> , 1982). Score 1-5 pour les 2 sous-échelles	NAP-sport NAP-loisirs	P-B	R

CONCEPT	QUESTIONNAIRE	VARIABLE (abréviation)	USAGE*	
Sentiment d'efficacité personnel pour l'exercice (SEPE)	L'échelle d'auto-efficacité des obstacles (<i>Barriers Self-Efficacy Scale</i> – 5 min) répertorie 13 raisons qui peuvent faire barrière à faire de l'exercice (McAuley, 1992). L'échelle de réponse varie de 0 % (pas du tout confiant) à 100 % (très confiant) et le score total (score / 100) représente la moyenne des 13 items. Cohérence interne : alpha de Cronbach de 0,92 (McAuley <i>et al.</i> , 2003).	SEPE	P-B	R
Soutien social/familial à l'exercice	Items (n = 13/50) de l'échelle <i>Friend/Peer Support-Health Eating Physical Activity Scale</i> (Sallis <i>et al.</i> , 1987). Échelles famille et ami, avec toutes deux les sous-échelles Récompense/Punition (RP) et Participation (Pa), mais l'échelle ami RP ne s'est pas avérée valide. Score/100	SSFEfamRP SSFEfamPa SSFEamiPa	P-B	R
Perception de la maladie	<i>Brief Illness Perception Questionnaire</i> (Brief-IPQ) – 8 items (5 min.) (Broadbent <i>et al.</i> , 2006). Score/80	PercMaladie	P-B	R
Outil de dépistage	STarT Back– 9 items (Hill <i>et al.</i> , 2008). Score total/9 et sous-score/5 (sous-échelle psychologique) Risque faible : score total < 4; risque moyen : score total ≥ 4 et sous-score ≤ 3; risque élevé : score total ≥ 4 et sous-score ≥ 4	STarT Back: 0 si « faible risque » et 1 si « moyen ou haut risque »	P-C	R
Attentes de traitement	<i>Outcome expectation for exercise scale</i> (OEES) – 9 items (4 min) (Resnick <i>et al.</i> , 2000). Score 1-5	AttentesTrait	P-B	R
Crédibilité/Attentes du traitement	<i>Credibility/Expectancy Questionnaire</i> (CEQ) – 6 items (3 min) (Deville et Borkovec, 2000). Score 11-100 pour les 2 sous-échelles	CEQ-crédibilité CEQ-attentes	/	R
Alliance thérapeutique patient/thérapeute	<i>Work Alliance Inventory</i> (WAI) – 36 items (5 min.) (Corbière <i>et al.</i> , 2006; Horvath et Greenberg, 1989). Score 1-7	AllianceTrav	/	R
Stades de changement du niveau d'activité physique	Questionnaire sur les stades de changement – 4 items (2 min) (Marcus <i>et al.</i> , 1992). Score 1-5 : <i>precontemplators</i> (1), <i>contemplators</i> (2); <i>preparers</i> (3), <i>actors</i> (4) and <i>maintainers</i> (5)	StadeChange-D2 est la variable dichotomisée : 0 si score = 1 1 si score = 2, 3, 4 ou 5	P-B	R
Adhésion aux exercices maison	Estimation (à T8 et T34) de l'adhésion aux exercices à domicile. Ratio entre 0 et 1 (fréquence réalisée / fréquence prescrite)	Adhésion	/	R

* Description du rôle joué par chaque variable dans l'analyse qui sera menée pour répondre à chaque objectif : Objectif 1 - Développement des règles de prédiction clinique (RPC); Objectif 2 - Détermination des mécanismes d'action associés au programme d'exercice.

G : variables utilisées pour créer des sous-groupes au regard du succès ou de l'échec du traitement.

P-A, P-B, P-C : prédicteurs (candidats aux RPC de succès) classés en classe A (lien direct ou indirect avec la stabilité lombaire), classe B (potentiellement liés à l'adhésion aux exercices à domicile) et les variables de classe C (non spécifiques; potentiellement liées à tout programme d'exercice). Voir la section 4.5.1 (analyse statistique) pour plus de détails (modélisation hiérarchique).

R : mesure de résultats (*outcome measure*). Pour l'objectif 2, les autres mesures de résultats ont été obtenues à partir des évaluations en laboratoire (mesures neuromusculaires).

Tableau A-2. Mesures de l'examen clinique de physiothérapie (PHT)

Examen clinique de physiothérapie (mesures PHT) de 60 min, présenté en ordre chronologique. Il faut savoir que les questionnaires (mesures OUT et PSY) étaient introduits à différents moments afin de reposer le patient entre les tests cliniques.

CONCEPT	INSTRUMENT OU TEST	VARIABLE (abréviation)	USAGE†	
			Obj.1	Obj.2
Mobilité chaîne postérieure	Test passif d'élévation jambe tendue selon Waddell (1997) avec inclinomètre sur épine tibiale. Mesure en degrés (EDM – étendue de mouvement) lors de l'apparition de douleur (peu importe la source) et en élévation maximale. Fidélité interjuges : $k = 0,67$ à $0,96$ (Bertilson <i>et al.</i> , 2006; McCarthy <i>et al.</i> , 2007; Waddell <i>et al.</i> , 1992)	EJT-Pas-Doul-EDM (d et g) EJT-Pas-Max-EDM (d et g)	P-C	R
Transfert de charge lombo-pelvienne	Test actif (20 cm) d'élévation jambe tendue selon Mens <i>et al.</i> (2001) : score sur 5 points avec niveau de difficulté observé par le thérapeute. Ensuite, le patient devait dire s'il ressentait ou non de la douleur. Fidélité interjuges : $k = 0,70$ à $0,77$ (Mens <i>et al.</i> , 2001; Roussel <i>et al.</i> , 2007; Waddell <i>et al.</i> , 1992)	EJT-Act/5 (d et g) EJT-Act-Doul (d et g)	P-A	R
Courbure lombaire	Avec un ou deux inclinomètres (Ng <i>et al.</i> , 2001). Fidélité intrajuge (interval : 3 jours): CCIC = $0,95$ (Ng <i>et al.</i> , 2001). Fidélité interjuges : CCIC = $0,97$ (Waddell <i>et al.</i> , 1992). Dans la présente étude, deux inclinomètres ont été utilisés.	Courbure	P-A	R
Étendue de mouvement lombaire (flexions avant) et pelvien (flexion) en position debout	Avec un (T12/L1) ou deux (T12/L1 et Sacrum) inclinomètres, selon Saur <i>et al.</i> (1996). Fidélité intrajuge (interval : 3 jours): CCIC = $0,87$ (Ng <i>et al.</i> , 2001). Fidélité interjuges : CCIC = $0,87$ à $0,89$ (Saur <i>et al.</i> , 1996; Waddell <i>et al.</i> , 1992). Dans la présente étude, deux inclinomètres ont été utilisés.	FlxLomb-EDM FlxBassin-EDM	P-C	R
Étendue de mouvement lombaire en flexion latérale en position debout	Avec un inclinomètre, selon Waddell <i>et al.</i> (1992). Fidélité intrajuge (interval : 3 jours): CCIC = $0,92$ à $0,96$ (Ng <i>et al.</i> , 2001). Fidélité interjuges : CCIC = $0,92$ à $0,95$ (Waddell <i>et al.</i> , 1992).	FlxLatLomb-EDM (d et g)	P-C	R
Questionnaires (n = 5)				
Instabilité lombaire	Test manuel d'instabilité lombaire, couché sur le ventre sur le rebord de la table d'examen, selon Stanton <i>et al.</i> (2011). Fidélité interjuges : $k = 0,67$ à $0,87$ (Alyazedi <i>et al.</i> , 2015; Fritz <i>et al.</i> , 2005; Hicks <i>et al.</i> , 2003; Puntumetakul <i>et al.</i> , 2015; Rabin <i>et al.</i> , 2013), mais des résultats moins élevés ($k = 0,27$ à $0,54$) ont aussi été observés (Ravenna <i>et al.</i> , 2011; Schneider <i>et al.</i> , 2008).	Instabilité-Man	P-A	R

CONCEPT	INSTRUMENT OU TEST	VARIABLE (abréviation)	USAGE†	
			Obj.1	Obj.2
Tests de déficience du contrôle sensorimoteur (DCM)	Symptômes (augmentation [1], diminution [-1], stables [0]) lors de mouvements passifs (n = 12) et actifs (n = 14), par rapport à une position de référence (k : 0,87 à 1,0) et étendue de mouvement (EDM)* au moment de l'apparition de la douleur (Van Dillen <i>et al.</i>, 2001; Van Dillen <i>et al.</i>, 2003): (1) Flexion genou (côtés droit et gauche) (2) Rotation de la hanche droite (rotations interne et externe) (3) Rotation de la hanche gauche (rotations interne et externe) (4) Extension 10° hanche (côtés droit et gauche) (5) Flexion épaule (côtés droit et gauche) – en mode actif seulement (6) Extension genou (côtés droit et gauche) (7) Abduction et rotation latérale hanche (côtés droit et gauche) Fidélité interjuges : k = 0,87 à 1,00 (Van Dillen <i>et al.</i>, 1998). <u>Réduction et dichotomisation des résultats*</u> : Les valeurs maximales (max) des 2 côtés ont été calculées. Ensuite, les variables (valeurs de -1, 0 ou 1) ont été dichotomisées : 0 si -1 ou 0 et 1 si 1 * L'EDM n'a été mesurée que pour les rotations internes et externes de la hanche lors des mouvements passifs.	(1) Flexion du genou GenFlx-Pas (d et g) GenFlx-Act (d et g) (2-3) Rotation de la hanche HanRI-Pas-EDM (d et g) † HanRE-Pas-EDM (d et g) † HanRI-Pas (d et g) HanRE-Pas (d et g) HanRI-Act (d et g) HanRE-Act (d et g) (4) Extension de la hanche HanExt-Pas (d et g) HanExt -Act (d et g) (5) Épaule ÉpaFlx-Act (d et g) (6) Extension du genou GenExt-Pas (d et g) GenExt-Act (d et g) (7) Abd. + rot. de la hanche HanRot-Pas (d et g) HanRot-Act (d et g)	Obj.1 P-A Sinon P-C †	Obj.2 R

* La relation, s'il elle existe, entre les tests DCM et l'instabilité lombaire est inconnue. C'est la raison pour laquelle ces différentes stratégies de réduction et de dichotomisation des résultats ont été évaluées.

CONCEPT	INSTRUMENT OU TEST	VARIABLE (abréviation)	USAGE†	
			Obj.1	Obj.2
Mouvements aberrants	Observations de 5 signes lors de la flexion/extension du tronc, en position debout ($k = 0,60 - 0,65$) (Hicks <i>et al.</i> , 2003); Score 0 (0 signe) ou 1 (1 signe + ou plus). Fidélité interjuges : $k = 0,60$ à $0,97$ (Alyazedi <i>et al.</i> , 2015; Hicks <i>et al.</i> , 2003; Puntumetakul <i>et al.</i> , 2015; Rabin <i>et al.</i> , 2013), mais des résultats moins élevés ($k = -0,07$) ont aussi été observés (Fritz <i>et al.</i> , 2005).	Mvt-Abe	P-A	R
Questionnaires (n = 4)				
Instabilité lombaire	Échelle des signes d'instabilité lombaire (15 questions) selon (Cook <i>et al.</i> 2006). Avec échelle de Likert de 5 points (pas du tout d'accord [1] à tout à fait d'accord [5]); Score / 75	Instabilité-15Q	/	R
Tests de performance physique (TPP)	Selon Novy <i>et al.</i> (2002): (1) 5 assis-debout répétés (temps); (2) 5 flexions répétées du tronc (temps); (3) portée avec charge (5 % poids) dans les mains, divisée par longueur du bras (ratio); (4) Rouler 360° sur matelas au sol à droite (temps) et à gauche (temps). Fidélité interjuges : CCIC = $0,81$ à $0,99$ (Novy <i>et al.</i> , 2002; Simmonds <i>et al.</i> , 1998)	TPP-assis TPP-flexions TPP-portée TPP-rouler (d et g)	P-A	R
Questionnaires (n = 8)				
Tests d'endurance musculaire (TEM)	Selon McGill <i>et al.</i> (1999) et Ito <i>et al.</i> (1996): Temps de maintien de quatre positions statiques permettant de mettre à contribution les différents groupes musculaires du tronc : muscles latéraux, abdominaux et dorsaux. Fidélité intrajuge (interval : 5 jours ou 8 semaines): CCIC = $0,93$ à $0,99$ (Ito <i>et al.</i> , 1996; McGill <i>et al.</i> , 1999).	TEM-lat (d et g) TEM-abdominaux TEM-dorsaux	P-A	R

CCIC : coefficient corrélation intraclass; k = coefficient kappa

(d et g) : signifie que la mesure a été prise pour le côté droit et ensuite pour le côté gauche

EDM : étendue de mouvement

† Description du rôle joué par chaque variable dans l'analyse qui sera menée pour répondre à chaque objectif : Objectif 1 - Développement des règles de prédiction clinique (RPC); Objectif 2 - Détermination des mécanismes d'action associés au programme d'exercice

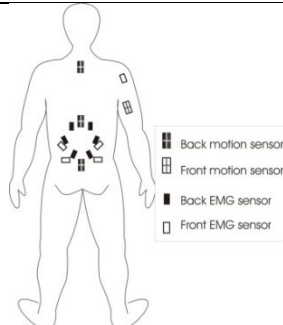
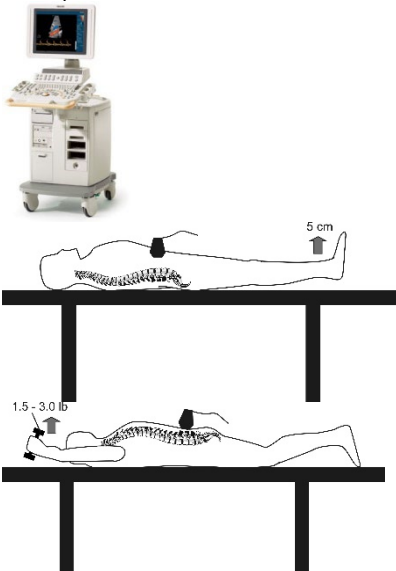
G : variables utilisées pour créer des sous-groupes au regard du succès ou de l'échec du traitement.

P-A, P-B, P-C: prédictors (candidats aux RPC de succès) classés en classe A (lien direct ou indirect avec la stabilité lombaire), classe B (potentiellement liés à l'adhésion aux exercices à domicile) et les variables de classe C (non spécifiques; potentiellement liées à tout programme d'exercice). Voir la section 4.5.1 (analyse statistique) pour plus de détails (modélisation hiérarchique).

R : mesure de résultats (*outcome measure*). Pour l'Objectif 2, les autres mesures de résultats ont été obtenues à partir des évaluations en laboratoire (mesures neuromusculaires), telles que présentées à l'**annexe A.3**.

Annexe A.3. Mesures neuromusculaires (NRM) en laboratoire

Ces tests évaluent le niveau de contrôle des muscles dorsaux et abdominaux. La procédure complète (préparation et 6 tests) durait environ trois heures.

I – Préparation du sujet : • Mesures anthropométriques et préparation de la peau; • Positionnement des senseurs de mouvement et des électrodes sur la peau et tâches de calibrage; • Six électrodes sur les muscles du dos; 4 électrodes sur les muscles abdominaux et 1 électrode sur l'épaule droite • Quatre senseurs de mouvement (tronc et bras droit) • Calibrage des senseurs à l'aide de postures statiques (n=4) et de contractions musculaires sous-maximales (n=6)	
II- Évaluation neuromusculaire – 6 tests (120 min, incluant familiarisation et périodes de repos) La séquence a été déterminée en fonction des capteurs à ajouter ou à retirer pour chaque test.	
Test n° 1 : Épaisseur et activation des muscles profonds (Larivière <i>et al.</i>, 2018; Larivière, Henry, <i>et al.</i>, 2019)  Tâche : levée de 5 cm de la jambe, en position couchée sur le dos et puis sur le ventre; 6 essais au total avec repos 1 min entre les essais. Mesure : l'épaisseur des muscles du tronc est quantifiée sur les images prises au repos et à la fin de la contraction.	Test n° 2 : Proprioception lombaire (Boucher <i>et al.</i>, 2017)  Tâche : les yeux fermés, un moteur fait très lentement tourner la chaise (2,2°/s) sur un côté jusqu'à l'atteinte de 10° (position cible). Le sujet a 3 sec. pour mémoriser cette position. Le moteur replace ensuite la chaise à 0°. Le sujet devait finalement retourner vers la position cible, mais cette fois-ci en utilisant ses muscles du tronc jusqu'à la position mémorisée, et presser le bouton-poussoir afin d'identifier la position finale; 10 essais consécutifs, soit 5 de chaque côté. Mesure : un potentiomètre mesure l'angle de la chaise lorsque le bouton est activé. La différence entre la cible de 10° et la position atteinte est la mesure.
	Test n° 3 : Équilibre postural en position assise sur chaise instable (Shahvarpour <i>et al.</i>, 2018)  Tâche : les yeux fermés, l'équilibre doit être maintenu pendant 60 s sur une chaise instable installée sur un pivot et des ressorts; 3 essais de pratique et 3 autres essais, avec un repos de 2 min entre les essais. Mesure : un senseur mesure les mouvements angulaires de la chaise.

Annexe A.3. Mesures neuromusculaires (NRM) en laboratoire (suite)

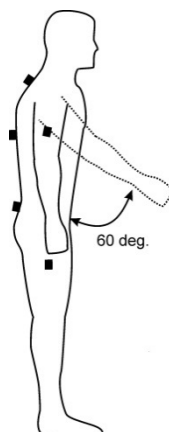
Test n° 4 : Rigidité lombaire (Larivière *et al.*, 2015)



Tâche : le sujet est immobilisé en position debout dans cet appareil, avec le bassin stabilisé et le thorax emprisonné dans un harnais rigide. Un essai consiste en une série de 125 perturbations soudaines du tronc (4 mm avant/arrière), ce qui produit une sensation de vibration. 5 essais de 75 sec avec repos de 2 min entre les essais.

Mesures : une cellule de force et un senseur de mouvement sont attachés entre le moteur et le harnais.

Test n° 5 : Ajustements posturaux anticipatoires (Boucher *et al.*, 2018)



Tâche : en réponse à un signal visuel affiché sur un écran positionné devant le sujet, celui-ci doit fléchir le bras droit vers l'avant (environ 60°) et retourner à la position de départ le plus rapidement possible. L'accent est placé sur la rapidité du mouvement plutôt que sur l'amplitude de 60°. 10 essais avec repos de 30 s entre les essais.

Mesures : activation musculaire (EMG de surface) de 10 muscles du tronc et mouvements du bras et de la région lombaire (senseurs de mouvement positionnés sur le sacrum [bassin]) et à T12].

Test n° 6 : Coordination du tronc (Shahvarpour *et al.*, 2017)



Tâche : en position debout avec les jambes droites, le sujet exécute 5 flexions maximales du tronc vers l'avant et retour en position debout (5 cycles).

Une cadence lente est imposée par un métronome pour chaque cycle de mouvements (flexion en 4 s + 2 s en flexion complète + 4 s retour en position debout).

Mesures : activation musculaire (EMG de surface) de 6 muscles dorsaux et mouvements des trois segments du tronc (bassin, colonnes lombaire et thoracique).

ANNEXE B. Résultats (objectif 1 : dérivation des RPC)

Tableau B-1. Comparaisons entre les sous-groupes succès (n = 54) et échec (n = 45) au temps T0 en ce qui a trait aux variables continues (qui ont été dichotomisées) – Résultats des participants utilisés pour la RPC à T8.

Variables*, seuil (méthode dichotomisation†)	Khi carré de Pearson	Indicateurs de performance diagnostique			
	Valeur $P \pm$	Se	Sp	LR+	LR-
Variables de classe A					
Beighton < 5/9 (KL)	0,115	93	18	1,13	0,42
Courbure < 24,5° (IY)	0,034	37	82	2,08	0,77
Courbure < 29° (Δ MinSnSp)	0,091	59	58	1,40	0,71
IMC $\geq$ 29,37 kg/m ² (ROC)	0,080	28	87	2,08	0,83
IMC $\geq$ 33,3 kg/m ² (LRmax)	0,037 ‡	15	98	6,67	0,87
TEM-abdominaux $\geq$ 71,9 s (ROC)	0,043	33	84	2,14	0,79
TEM-dorsaux $\geq$ 225,56 s (LRmax)	0,088	15	96	3,33	0,89
TPP-assis $\geq$ 7,6 s (LRmax)	0,044 ‡	98	13	1,13	0,14
TPP-flexions $\geq$ 7,48 s (ROC)	0,050	70	49	1,38	0,61
TPP-flexions $\geq$ 5,97 s (KLg/f)	0,075‡	96	16	1,14	0,24
TPP-portée $\geq$ 0,75/1 (ROC)	0,007	28	93	4,17	0,77
TPP-portée $\geq$ 0,8/1 (LRmax)	0,068‡	13	98	5,83	0,89
Variables de classe B					
AttentesTrait < 3,72 /5 (KLg/f)	0,177	24	87	1,81	0,88
FABQ- AP $\geq$ 6,5/24 (LRmax)	0,199	93	16	1,10	0,48
NAP-loisirs $\geq$ 1,62 /5 (ROC)	0,111	98	9	1,08	0,21
PercMaladie $\leq$ 44 /80 (ROC)	0,052	57	62	1,52	0,68
PercMaladie $\leq$ 38,5 /80 (KLf/g)	0,033	31	87	2,36	0,79
PDlang < 54,17 /100 (KLg/f)	0,075‡	96	16	1,14	0,24
PDldep < 2,8 /100 (KLg/f)	0,080	33	82	1,88	0,81
SSESfamPa < 34,5 /100 (KLf/g)	0,019	93	24	1,23	0,30
Variables de classe C					
FlxBassin-EDM < 73° (KLg/f)	0,095	50	67	1,50	0,75
FlxLatLomb-EDM-Min < 19° (KLg/f)	0,005	44	82	2,50	0,68
FlxLomb-EDM $\geq$ -3° (ROC)	0,051	20	93	3,06	0,85
HanRE-Pas-EDM-Min < 61° (KLf/g)	0,008	67	60	1,67	0,56
HanRI-Pas-EDM-Min < 51° (KLf/g)	0,022 ‡	98	16	1,16	0,12
EJT-Pas-Doul-EDM-Min < 67° (KLf/g)	0,090	83	31	1,21	0,54
EJT-Pas-Max-EDM-Moy < 92,5° (KLf/g)	0,135	89	22	1,14	0,50
EJT-Pas-Max-EDM-Min < 51° (KLf/g)	0,106‡	15	96	3,33	0,89
EJT-Pas-Max-EDM-Min < 97° (KLg/f)	0,089‡	98	11	1,10	0,17
Int-Douleur $\leq$ 6,5 / 10 (KLg/f)	0,014	94	22	1,21	0,25

* Seulement les variables qui ont démontré une tendance ($P < 0,20$) avec le test Khi carré ou ayant un LR+ > 2 ont été retenues comme candidates et présentées dans ce tableau, à l'exception des variables d'ajustement du modèle (âge, IMC). Les variables retenues dans les modèles sont en caractères gras.

† Méthode de dichotomisation : IY : index de Youden; KL : distances de Kullback-Leibler; LRmax = maximisation du rapport de vraisemblance (LR); Med : médiane; ROC = point optimum courbe ROC; Δ MinSnSp : différence minimale entre le taux des vrais positifs et le taux des vrais négatifs.

‡ Test exact de Fisher pour les effectifs faibles.

Tableau B-2. Comparaisons entre les sous-groupes succès (n = 54) et échec (n = 45) au temps T0 en ce qui a trait aux variables dichotomiques – Résultats des participants utilisés pour la RPC à T8.

Variables*	Khi carré de Pearson	Indicateurs de performance diagnostique			
	Valeur <i>P</i> ‡	Se	Sp	LR+	LR-
Variables de classe A					
Sexe féminin	0,397	57	51	1,17	0,83
EJT-Act-Doul-Max positif	0,013	46	78	2,08	0,69
Instabilité-Man négatif	0,042	65	56	1,46	0,63
Mvt-Abe positif	0,022	39	82	2,19	0,74
HanRE-Act-Max positif	0,191	46	67	1,39	0,81
ÉpaFlx-Act-Max positif	0,138	33	80	1,67	0,83
GenExt-Pas-Max positif	0,095	50	67	1,50	0,75
GenExt-Act-Max positif	0,136	48	67	1,44	0,78
HanRot-Pas-Max positif	0,004	48	80	2,41	0,65
HanRot-Act-Max positif	0,042	39	80	1,94	0,76
Variables de classe B					
StadeChange-D2 positif	0,042‡	100	9	1,10	0,00
Variables de classe C					
Aucune					

* Seulement les variables qui ont démontré une tendance ($P < 0,20$) avec le test Khi carré ou ayant un LR+ > 2 ont été retenues comme candidates et présentées dans ce tableau, à l'exception des variables d'ajustement du modèle (sexe).

Les variables qui ont ultérieurement été retenues dans les différents modèles multivariés (RPC de succès) sont en caractères gras.

‡ Test de Pearson et test exact de Fisher (‡) pour les effectifs faibles. Les valeurs $P \leq 0,05$ sont en caractères gras.

Tableau B-3. Comparaisons entre les sous-groupes succès (n = 53) et échec (n = 36) au temps T0 en ce qui a trait aux variables continues (qui ont été dichotomisées) – Résultats des participants utilisés pour la RPC à T34.

Variables*, seuil (méthode dichotomisation†)	Khi carré de Pearson Valeur P ‡	Indicateurs de performance diagnostique (cellules ombragées ‡‡)			
		Se	Sp	LR+	LR-
Variables de classe A					
Âge < 43 ans (Md)	0,547	51	56	1,15	0,88
Beighton > 0,0/9	0,163	45	69	1,48	0,79
IMC ≥ 33,65 (LRmax)	0,644	6	97	2,04	0,97
TEM-abdominaux ≥ 72,62 s (KLf/g)	0,022	36	86	2,58	0,74
TEM-dorsaux ≥ 236,18 s (KLf/g)	0,078‡	15	97	5,43	0,87
TPP-assis ≥ 11,2 s (LRmax)	0,100	51	67	1,53	0,74
TPP-assis ≥ 7,2 s (KLf/g)	0,063‡	99	8	1,08	0,11
TPP-flexions ≥ 9,91 s (IY)	0,071	34	83	2,04	0,79
TPP-flexions ≥ 11,4 s (KLf/g)	0,078	26	89	2,38	0,83
TPP-portée ≥ 0,76/1 (KLf/g)	0,012	26	94	4,75	0,78
TPP-portée ≥ 0,67/1 (ROC)	0,040	43	78	1,95	0,73
TPP-portée ≥ 0,63/1 (ΔMinSnSp)	0,119	58	58	1,40	0,71
TPP-portée ≥ 0,80/1 (LRmax)	0,019‡	15	99	11,02	0,86
Variables de classe B					
AttentesTrait ≥ 4,22 /5 (Md)	0,069	58	61	1,50	0,68
FABQ- AP < 12/24 (KLf/g)	0,029	42	81	2,13	0,73
Dramatisation/douleur < 23/52 (KLf/g)	0,113	64	53	1,36	0,68
NAP-loisirs ≥ 1,75 /5 (LRmax)	0,153‡	98	11	1,10	0,17
PDIsog < 25 /100 (ΔMinSnSp)	0,076	43	75	1,74	0,75
PDIsoma < 9,1 /100 (LRmax)	0,078‡	15	97	5,43	0,87
PercMaladie < 44 /80 (Md)	0,025	55	69	1,79	0,65
PercMaladie < 45 /80 (KLf/g)	0,007	62	67	1,87	0,57
Variables de classe C					
EJT-Pas-Doul-EDM-Min < 68° (KLf/g)	0,022	85	36	1,33	0,42
EJT-Pas-Max-EDM-Moy < 93° (KLf/g)	0,155	87	25	1,16	0,53
EJT-Pas-Max-EDM-Min < 52° (KLf/g)	0,044‡	17	97	6,11	0,85
FlxLatLomb-EDM-Min ≥ 36° (LRmax)	0,234‡	11	97	4,08	0,91
FlxLomb-EDM < -16° (ΔMinSnSp)	0,058	57	64	1,57	0,68
FlxLomb-EDM ≥ -2° (ROC)	0,188‡	17	94	3,06	0,88

* Seulement les variables qui ont démontré une tendance ($P < 0,20$) avec le test Khi carré ou ayant un $LR+ > 2$ ont été retenues comme candidates et présentées dans ce tableau, à l'exception des variables d'ajustement du modèle (âge, IMC). Les variables retenues dans les modèles sont en caractères gras.

† Méthode de dichotomisation : IY : index de Youden; KL : distances de Kullback-Leibler; LRmax = maximisation du rapport de vraisemblance (LR); Med : médiane; ROC = point optimum courbe ROC; ΔMinSnSp : différence minimale entre le taux des vrais positifs et le taux des vrais négatifs.

‡ Test exact de Fisher pour les effectifs faibles.

‡‡ Lorsqu'une cellule du tableau de contingence était vide (0), 0,5 était ajouté pour permettre l'estimation de ces statistiques (à interpréter avec prudence).

Tableau B-4. Comparaisons entre les sous-groupes succès (n = 53) et échec (n = 36) au temps T0 en ce qui a trait aux variables dichotomiques – Résultats des participants utilisés pour la RPC à T34.

Variables*	Khi carré de Pearson	Indicateurs de performance diagnostique			
	Valeur P ‡	Se	Sp	LR+	LR-
Variables de classe A					
Sexe masculin	0,800	47	56	1,06	0,95
HanRI-Pas-Max positif	0,196	53	61	1,36	0,77
HanExt-Act-Max négatif	0,109	42	75	1,66	0,78
Mvt-Abe positif	0,044	40	81	2,04	0,75
Variables de classe B					
Aucune					
Variables de classe C					
StartBack (moyen ou haut risque)	0,066	47	72	1,70	0,73

* Seulement les variables qui ont démontré une tendance ($P < 0,20$) avec le test Khi carré ou ayant un $LR+ > 2$ ont été retenues comme candidates et présentées dans ce tableau, à l'exception des variables d'ajustement du modèle (sexe).

Les variables qui ont ultérieurement été retenues dans les différents modèles multivariés (RPC de succès) sont en caractères gras.

‡ Test de Pearson. Les valeurs $P \leq 0,05$ sont en caractères gras alors que les valeurs situées entre 0,05 et 0,20 sont soulignées.