

La version franco-canadienne de l'outil *Assessment of Multiple Systematic Reviews* (AMSTAR)

Lucie Brosseau, Ph. D.;* Chantal Laroche, Ph. D.;† Paulette Guitard, Ph. D.;‡ Judy King, Ph. D.;* Stéphane Poitras, Ph. D.;* Lynn Casimiro, Ph. D.;§ Julie Alexandra Barette, MLing;¶ Dominique Cardinal, M.A.;** Sabrina Cavallo, Ph. D.;†† Lucie Laferrière, MHA, PT;‡‡ Rose Martini, Ph. D.;* Nicholas Champoux, MScPT;* Jennifer Taverner, MScPT;* Chanyque Paquette, MScPT;* Sébastien Tremblay, MScPT*

COMITÉ D'EXPERTS POUR LA RÉDACTION DE LA PREMIÈRE VERSION EXPÉRIMENTALE DE L'OUTIL AMSTAR

Ann Sutton, Ph. D.;† Roseline Galipeau, Ph. D.;§§ Jocelyne Tourigny, Ph. D.;¶¶ Karine Toupin-April, Ph. D.;*** Laurianne Loew, MSc, Ph. D. (cand.);†† Catrine Demers, MSc, Ph. D. (cand.);†† Katrine Sauvé-Schenk, MPH, Ph. D. (cand.)††

COMITÉ D'EXPERTS POUR L'ÉVALUATION DE LA DEUXIÈME VERSION EXPÉRIMENTALE DE L'OUTIL AMSTAR

Nicole Paquet, Ph. D.;* Jacinthe Savard, Ph. D.;‡ José Lagacé, Ph. D.;† Denyse Pharand, Ph. D.;¶¶ Véronique Vaillancourt, MSc†

RÉSUMÉ

Objectif : le premier but est de produire une traduction franco-canadienne de l'outil AMSTAR et d'examiner la validité de son contenu. Les second et troisième buts sont d'examiner la fiabilité interjuges et la validité de construit factorielle de la version franco-canadienne de l'outil AMSTAR. **Méthodes :** une approche modifiée de la méthodologie de validation transculturelle de Vallerand (1989) a été adoptée.¹ Une traduction renversée parallèle de l'outil AMSTAR² a d'abord été effectuée à la fois par des professionnelles et de futurs professionnels. Ensuite, un premier comité d'experts (P1) a examiné les versions traduites et a créé l'ébauche expérimentale préliminaire de l'outil AMSTAR. Celle-ci a été évaluée et modifiée par un deuxième comité d'experts (P2). Finalement, 18 futurs professionnels à la maîtrise en physiothérapie ont évalué cette deuxième version expérimentale de l'outil à l'aide d'une échelle d'ambiguïté de 7 points (1 : très clair; 7 : très ambigu). Les cochercheurs principaux ont examiné les éléments problématiques et ont proposé les modifications finales. Vingt revues systématiques publiées en français après l'an 2000 ont été évaluées par quatre juges indépendants à l'aide de la version franco-canadienne de l'outil AMSTAR. Le coefficient de corrélation intraclasse (CCI) et le coefficient kappa ont été utilisés afin d'examiner sa fiabilité interjuges. Le coefficient alpha Cronbach a été calculé pour évaluer la cohérence interne. L'analyse factorielle a été utilisée pour examiner la validité de construit afin de déterminer le nombre de dimensions. **Résultats :** pour les différents énoncés de la version finale de l'outil AMSTAR, les moyennes sur l'échelle d'ambiguïté varient entre 1,0 et 1,4. Aucun énoncé n'a obtenu une moyenne en deçà de 1,4 démontrant ainsi un niveau élevé de clarté. La fiabilité interjuges ($n = 4$) pour le score total de l'outil AMSTAR est qualifiée de modérée selon le coefficient de corrélation intraclasse d'une valeur de 0,61 (intervalle de confiance [IC] à 95 % : 0,29, 0,97). La fiabilité interjuges est considérée bonne pour 82 % les éléments individuels selon les valeurs de kappa obtenues. La cohérence interne est excellente d'après le coefficient alpha de Cronbach représente une valeur de 0,91 (IC à 95 % : 0,83, 0,99). La version franco-canadienne de l'outil AMSTAR est un outil unidimensionnel confirmé par l'analyse factorielle et les valeurs de communauté

Affiliations : *Programme de physiothérapie; †Programmes d'audiologie et d'orthophonie; ‡Programme d'ergothérapie, École des sciences de la réadaptation; ¶Département de linguistique; **Consortium national de formation en santé (CNFS); ††École des sciences de la réadaptation; ¶¶École des sciences infirmières, Université d'Ottawa; §Hôpital Montfort; ***Institut de recherche de l'hôpital pour enfants de l'Est de l'Ontario; ‡‡Direction de la protection de la santé des Forces armées canadiennes, Ottawa; §§Département des sciences infirmières, Université du Québec en Outaouais, Gatineau, Qué.

Correspondance : Lucie Brosseau (Ph. D.), Professeure titulaire, Faculté des sciences de la santé, École des sciences de la réadaptation, Université d'Ottawa, 451 Smyth, Pièce : RGN3060, Ottawa, ON K1H 8M5 Canada; Lucie.Brosseau@uottawa.ca.

Collaborateurs : Tous les auteurs ont conçu l'étude; ou ont recueilli, analysé ou interprété les données; et ont rédigé ou procédé à l'examen critique de l'article et en ont approuvé la version définitive.

Conflits d'intérêts : Les auteurs n'ont aucun conflit d'intérêt à déclarer.

Remerciements : Les auteurs remercient le cabinet du Vice-rectorat associé aux études, la Chaire de recherche universitaire en réadaptation (LB) de l'Université d'Ottawa pour le soutien financier apporté à ce projet. Les auteurs sont très reconnaissants envers Mesdames Luce Ouellet et Claire Mazuhelli, traductrices professionnelles, ainsi que Madame Rachel Marcotte et Madame Prinon Rahman, assistantes de recherche pour leur précieuse contribution.

Physiotherapy Canada 2017; 69(1);20–29; doi:10.3138/ptc.2015-80F

supérieures à 0,30. **Conclusion :** les cinq étapes rigoureuses du processus ont permis de produire une version franco-canadienne valide de l'outil AMSTAR. La version franco-canadienne de l'outil AMSTAR est un outil unidimensionnel qui offre une fiabilité interjuges modérée pour l'ensemble des éléments et une excellente cohérence interne. Cet outil peut s'avérer pertinent pour les professionnels et chercheurs franco-canadiens et pourrait intéresser la francophonie internationale.

ABSTRACT

Objectives: The primary objective was to produce a French-Canadian translation of AMSTAR (a measurement tool to assess systematic reviews) and to examine the validity of the translation's contents. The secondary and tertiary objectives were to assess the inter-rater reliability and factorial construct validity of this French-Canadian version of AMSTAR. **Methods:** A modified approach to Vallerand's methodology (1989) for cross-cultural validation was used.¹ First, a parallel back-translation of AMSTAR² was performed, by both professionals and future professionals. Next, a first committee of experts (P1) examined the translations to create a first draft of the French-Canadian version of the AMSTAR tool. This draft was then evaluated and modified by a second committee of experts (P2). Following that, 18 future professionals (master's students in physiotherapy) rated this second draft of the instrument for clarity using a seven-point scale (1: very clear; 7: very ambiguous). Lastly, the principal co-investigators then reviewed the problematic elements and proposed final changes. Four independent raters used this French-Canadian version of AMSTAR to assess 20 systematic reviews that were published in French after the year 2000. An intraclass correlation coefficient (ICC) and kappa coefficient were calculated to measure the tool's inter-rater reliability. A Cronbach's alpha coefficient was also calculated to measure internal consistency. In addition, factor analysis was used to evaluate construct validity in order to determine the number of dimensions. **Results:** The statements on the final version of the AMSTAR tool received an average ambiguity rating of between 1.0 and 1.4. No statement received an average rating below 1.4, which indicates a high level of clarity. Inter-rater reliability ($n = 4$) for the instrument's total score was moderate, with an intraclass correlation coefficient of 0.61 (95% confidence interval [CI]: 0.29, 0.97). Inter-rater reliability for 82% of the individual items was good, according to the kappa values obtained. Internal consistency was excellent, with a Cronbach's alpha coefficient of 0.91 (95% CI: 0.83, 0.99). The French-Canadian version of AMSTAR is a unidimensional tool, as confirmed by factor analysis and community values greater than 0.30. **Conclusion:** A valid French-Canadian version of AMSTAR was created using this rigorous five-step process. This version is unidimensional, with moderate inter-rater reliability for the elements overall, and with excellent internal consistency. This tool could be valuable to French-Canadian professionals and researchers, and could also be of interest to the international Francophone community.

Key Words: research methodology; systematic reviews, as topic; translation, cross-cultural validity.

INTRODUCTION

Une revue systématique (RS) est une synthèse d'études scientifiques recueillies de façon rigoureuse et analysées de nouveau afin de faire ressortir un résultat synthèse portant le plus souvent sur l'efficacité d'une intervention en santé.¹ Afin de vérifier si les conclusions des études portant sur l'efficacité des interventions cliniques sont solides, le clinicien averti doit s'assurer de la bonne qualité méthodologique des écrits scientifiques.² Il existe des instruments de mesure validés en français de la qualité méthodologique pour les essais cliniques randomisés³ et pour les lignes directrices cliniques.⁴ Cependant, aucun instrument, conçu pour évaluer la qualité méthodologique des RS, n'existe en langue française. C'est pourquoi le but de ce projet est de produire une traduction franco-canadienne validée de l'outil AMSTAR, reconnu et utilisé mondialement.

L'outil AMSTAR («Assessment of Multiple Systematic Reviews»), est un instrument de mesure international et validé en langue anglaise qui sert à évaluer la qualité méthodologique des RS.⁵ L'outil AMSTAR est utile aux professionnels de la santé^{2,6-10} afin que ces derniers puissent juger de la qualité des RS et améliorer leurs décisions cliniques.

L'outil AMSTAR est constitué de 11 éléments méthodologiques propres aux RS: 1) la formulation de la question de recherche et des critères d'éligibilité; 2) le processus de sélection des études et d'extraction des données; 3) la recherche exhaustive des écrits scientifi-

ques; 4) l'identification du type de publication inclus; 5) la liste des études incluses et exclues; 6) la liste des caractéristiques des études incluses; 7) l'évaluation de la qualité scientifique des études incluses dans la RS; 8) l'évaluation de la qualité scientifique des études incluses dans la conclusion de la RS; 9) la pertinence des méthodes utilisées dans la synthèse des résultats; 10) la présence potentielle de biais; et 11) la spécification des conflits d'intérêt.⁵ Chaque élément de l'outil AMSTAR est évalué en cochant l'un de ces choix de réponses: «oui», «non», «ne peut répondre» ou «non applicable». Ce système de cotes vaut 1 point pour «oui» et 0 point pour les autres options de réponse. Le score total de l'outil AMSTAR correspond au nombre de réponses «oui» pour l'ensemble des éléments, soit un score total exprimé sur 11 points.

Plusieurs études ont examiné les qualités métrologiques de l'outil AMSTAR.^{5,7-9} Les résultats de l'étude de Shea et collaborateurs (2009),⁵ dans laquelle 30 RS furent évaluées par deux examinateurs, ont révélé que l'outil AMSTAR possède une bonne fiabilité interjuges (coefficient de corrélation intraclass [CCI] de 0,84 [IC à 95 %: 0,65, 0,92]). L'étude de fiabilité interjuges menée par Kang et collaborateurs (2012),⁸ dans laquelle 41 RS ont été évaluées, a également obtenu des résultats similaires (CCI élevé de 0,87; IC à 95 %: 0,76, 0,93). La fiabilité de l'outil AMSTAR n'est donc aucunement mise en doute.

Une étude de validité discriminante⁷ a été effectuée par l'analyse comparative de la qualité méthodologique,

Tableau 1 Caractéristiques des participants au processus de traduction

	Comité d'experts 1 (n = 7)	Comité d'experts 2 (n = 5)	Futurs professionnels de la santé (n = 18)
Âge	44,6 ± 12,3	47,1 ± 11,2	22,8 ± 1,5
Sexe	Féminin : 7 Masculin : 2	Féminin : 5	Féminin : 16
Langue maternelle	Français : 6 Anglais : 1	Français : 5	Français : 18
Profession	Audiologiste : 1 Ergothérapeutes : 2 Infirmières : 2 Orthophoniste : 1 Physiothérapeute : 1	Audiologiste : 1 Ergothérapeute : 1 Infirmière : 1 Orthophoniste : 1 Physiothérapeute : 1	Futur(e)s physiothérapeutes : 18
Années d'expérience en recherche (en moyenne)	8,9 ± 5,2	17,6 ± 7,3	1,2 ± 0,8

de 30 RS publiés en 2012. Ces RS étaient publiées soit : 1) dans des revues scientifiques qui exigeaient la liste de contrôle PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis*), lequel sert à rapporter et à identifier, par auteurs, tous les éléments constitutifs d'une RS (Groupe A) ou 2) celles qui n'exigeaient pas la liste de contrôle PRISMA (Groupe B). Aucune différence statistiquement significative ($p > 0,05$) n'a été observée entre la qualité méthodologique des RS, mesurée à l'aide de l'outil AMSTAR du groupe A (85,0 %, étendue 77,6–90,7) et celles du groupe B (85,6 % étendue 78,2–91,2). L'outil AMSTAR ne serait donc pas un outil discriminant.

Deux études de validité de construit factorielle^{5,9} ont examiné la structure dimensionnelle de l'outil AMSTAR. Ces études respectives ont permis de conclure que l'outil AMSTAR était unidimensionnel et que tous les 11 éléments constituant l'outil étaient nécessaires. À notre connaissance, bien que l'outil AMSTAR soit reconnu et très utilisé en tant qu'instrument fiable et validé en langue anglaise pour évaluer la qualité méthodologique des RS, aucune traduction validée française n'existe à l'heure actuelle. Ce projet vise à remédier à cette lacune et à proposer une version franco-canadienne valide de l'outil AMSTAR.

Le but du présent projet est de produire une traduction validée en langue franco-canadienne de l'outil AMSTAR par l'examen 1) de sa validité transculturelle concomitante et apparente; 2) de sa fiabilité interjuges; et 3) de sa validité de construit factorielle de la version franco-canadienne de l'outil AMSTAR.

MÉTHODOLOGIE

Comme le suggèrent Hébert, Bravo et Voyer (1994),¹¹ une version modifiée de la méthodologie de validation transculturelle de questionnaires conçue par Vallerand

(1989)¹² a été utilisée afin de valider l'équivalence franco-canadienne de l'outil AMSTAR.⁵ Conçue au Canada dans le domaine de la psychologie, cette méthodologie a été reprise par plusieurs autres disciplines et secteurs d'intervention dont la réadaptation.^{3,13–16} Cette étude a obtenu l'approbation déontologique (H12-13-03).

Cinq étapes systématiques ont été suivies pour traduire l'outil AMSTAR : 1) la traduction renversée parallèle (versions préliminaires); 2) la rédaction de la version expérimentale préliminaire; 3) l'évaluation de la première version expérimentale; et 4) l'évaluation de la deuxième version expérimentale et 5) la traduction inversée vers l'anglais de la version préofficielle afin de produire la version officielle franco-canadienne de l'outil AMSTAR. Deux comités ont été constitués et des participants ont été recrutés pour participer à ces différentes étapes. Les membres des comités et les participants consultés ont été invités à remplir un court questionnaire démographique, ainsi qu'un formulaire de consentement.

Étape 1 : traduction renversée parallèle

Cette étape préparatoire a permis de faire traduire l'outil AMSTAR en langue franco-canadienne par une première traductrice professionnelle (TP1) bilingue ayant comme langue maternelle le français (LO). Une traductrice clinicienne et chercheuse (TC1) bilingue (LL) et ayant comme langue maternelle le français en a fait également une traduction de l'anglais au franco-canadien. Ces deux versions franco-canadiennes ont par la suite été respectivement retraduites en anglais (TP2 et TC2) par une deuxième traductrice professionnelle (CM) et une deuxième clinicienne et chercheuse (SC), toutes deux bilingues et de langue maternelle anglaise. Les traductrices professionnelles (LO et CM) et cliniciennes (LL et SC) ont été recrutées par courriel et avaient déjà participé à d'autres validations transculturelles d'instruments^{3,15} (Tableau 1).

Étape 2 : élaboration de la version expérimentale à partir des versions préliminaires

Lors de cette étape, un comité d'experts composé de l'équipe de chercheurs (LB, NC, JT, CP, ST) et de sept professionnels de la santé bilingues œuvrant en recherche clinique dans la région d'Ottawa (SA, RG, JT, KTA, LL, CD, KSS) a examiné les cinq versions de l'outil AMSTAR (originale, TP1, TP2, TC1 et TC2) (Tableau 1). À partir de ces cinq versions, les membres du comité ont formulé une version expérimentale préliminaire de l'outil AMSTAR en français. Pour faire suite à la traduction, les membres du comité devaient d'abord lire et comparer la version originale en langue anglaise et la version traduite en langue française de l'outil. Ensuite, ils devaient vérifier si le sens des énoncés avait été conservé en s'assurant que les mots transmettaient la signification désirée. Si un mot ne concordait pas avec la signification voulue, le comité pouvait proposer un autre terme qui correspondait mieux au sens de la version originale anglaise de l'outil.

La version expérimentale préliminaire franco-canadienne devait respecter les trois consignes suivantes :

- tenter de dégager le sens de l'énoncé original et non le « mot pour mot » lorsqu'une traduction exacte n'était pas possible;
- accorder autant que possible une importance aux subtilités de la langue franco-canadienne dans la formulation;
- adopter le même format de présentation et les mêmes directives que ceux utilisés dans la version originale.

Les dictionnaires Petit Larousse et Petit Robert^{17,18} étaient disponibles sur place ainsi qu'un livre de méthodologie de la recherche en français.¹⁹

Étape 3 : prétest de la version expérimentale

Le but de cette étape était d'apporter, si nécessaire, des modifications à la version expérimentale élaborée à l'étape 2 à l'aide d'un deuxième comité d'experts. Celui-ci était composé de six professionnels de la santé bilingues œuvrant en recherche clinique (NP, JS, JL, DP, VV,) qui n'avaient pas encore participé au projet (Tableau 1). À cette étape, une trousse d'information a été préparée pour chacun des experts, comprenant les documents suivants :

- le questionnaire sociodémographique;
- la lettre de recrutement ainsi que le formulaire de consentement;
- l'outil original (« AMSTAR tool ») et la version expérimentale préliminaire de l'outil AMSTAR en langue franco-canadienne

Les dictionnaires Petit Larousse¹⁸ et Petit Robert¹⁹ étaient disponibles sur place ainsi que l'accès à un outil terminologique en ligne.²⁰

Les membres de ce second comité ont évalué la traduction de l'ensemble des éléments et modifié, s'il y avait lieu, certains éléments de la version expérimentale préliminaire après consensus. À la fin de cette étape, une première version expérimentale de l'outil AMSTAR en langue franco-canadienne a été produite.

Étape 4 : élaboration de la deuxième version expérimentale

À cette étape, 18 étudiants francophones en formation à la maîtrise professionnelle en physiothérapie ont évalué la première version expérimentale franco-canadienne de l'outil AMSTAR et ont proposé des modifications pour produire la deuxième version expérimentale. Des lettres d'invitation à participer au projet de recherche ont été envoyées à ces futurs professionnels en physiothérapie. Les chercheurs de la présente étude ont choisi cette population pour ces deux raisons principales : 1) les personnes qui reçoivent une formation francophone avaient déjà une bonne connaissance du vocabulaire de la réadaptation en français et de l'outil AMSTAR, puisqu'elles avaient utilisé une traduction maison de l'outil et 2) dans une étude similaire³ visant l'évaluation de la validité transculturelle concomitante et apparente de la traduction franco-canadienne d'un autre instrument de mesure en réadaptation, le taux de participation des professeurs-chercheurs affiliés aux programmes de réadaptation universitaires francophones du Canada avait été très bas (14 %). À la suite du recrutement des futurs professionnels de langue maternelle française en première année de physiothérapie ($n = 18$), ces derniers ont été convoqués à une séance d'évaluation par les chercheurs du projet (NC, JT, CP, ST), et des trousse d'information ont été fournies aux participants. Ces trousse comprenaient les documents suivants :

- le questionnaire sociodémographique;
- la lettre de recrutement ainsi que le formulaire de consentement;
- la première version expérimentale franco-canadienne de l'outil AMSTAR;
- une échelle d'ambiguïté de 7 points suggérée par la méthode Vallerand (1989)¹² (1 signifiant « En parfait accord » et 7 signifiant « En profond désaccord ») avec possibilité d'ajouter des commentaires et des suggestions.

À partir des commentaires reçus et des résultats des évaluations obtenus à l'aide de l'échelle d'ambiguïté, l'ensemble des chercheurs (LB, CL, PG, JK, SP, LC, JAB, DC, SC, LL, RN, NC, JT, CP, ST) a pu réviser la formulation des énoncés de la deuxième version expérimentale de l'outil et modifier, s'il y avait lieu, certains éléments, afin de créer la version préofficielle de l'outil AMSTAR. Ils ont procédé par consensus pour chacun des éléments afin de produire cette version, qui est décrite au Boîte 1.

Boîte 1 L'outil AMSTAR (encerclez un chiffre par ligne)

1. Est-ce qu'un devis d'étude a été établi a priori? La question de recherche et les critères d'inclusion devraient être établis avant de commencer la revue.	Oui Non Impossible de répondre Sans objet
2. La sélection des études et l'extraction des données ont-elles été effectuées à deux reprises? Au moins deux évaluateurs devraient extraire les données de manière indépendante et une procédure pour parvenir à un consensus devrait être établie afin de résoudre les désaccords.	Oui Non Impossible de répondre Sans objet
3. La recherche des écrits était-elle exhaustive? Au moins deux sources électroniques devraient être utilisées. Les années de publication et les bases de données utilisées doivent être indiquées dans le rapport (p. ex. Central Pubmed, EMBASE et MEDLINE, CINAHL). Les mots-clés ou les termes MESH doivent être mentionnés et, dans la mesure du possible, la stratégie de recherche devrait être décrite. La recherche devrait également comprendre la consultation de contenus à jour, de revues, de manuels, de répertoires spécialisés ou d'experts dans le domaine à l'étude, ainsi que l'examen de la bibliographie des études repérées.	Oui Non Impossible de répondre Sans objet
4. Est-ce que le type de publication (p.ex. littérature grise) a fait partie des critères d'inclusion? Les auteurs devraient indiquer s'ils ont recensé tous les écrits sans tenir compte du type de publication. Ces derniers devraient aussi indiquer s'ils ont exclu des rapports (de la revue systématique) en fonction du type de publication, de la langue, etc.	Oui Non Impossible de répondre Sans objet
5. Une liste des études retenues et rejetées a-t-elle été fournie? Une liste des études retenues et rejetées devrait être fournie.	Oui Non Impossible de répondre Sans objet
6. Les caractéristiques des études retenues ont-elles été fournies? Les données sur les participants, les interventions et les mesures de résultats provenant des études originales devraient être regroupées (p.ex. dans un tableau). L'étendue des valeurs liées aux caractéristiques rapportées dans les études analysées (p.ex., âge, race, sexe, données socioéconomiques pertinentes, état de santé, durée et sévérité de la maladie, présence d'autres maladies) devrait être présentée.	Oui Non Impossible de répondre Sans objet
7. La qualité scientifique des études retenues a-t-elle été évaluée et documentée? Les méthodes d'évaluation établies a priori devraient être indiquées. Par exemple, dans le cas d'études d'efficacité, les auteurs ont-ils indiqué s'ils ont choisi d'inclure uniquement les essais cliniques randomisés à double insu avec placebo, ou ceux dont l'attribution des traitements est dissimulée? Pour d'autres types d'études, différents éléments pourraient être pertinents.	Oui Non Impossible de répondre Sans objet
8. La qualité scientifique des études retenues a-t-elle été considérée de façon adéquate dans la formulation des conclusions? Les résultats de l'évaluation de la rigueur méthodologique et de la qualité scientifique devraient être pris en considération pour l'analyse et les conclusions de la revue, et explicitement énoncés dans la formulation des recommandations.	Oui Non Impossible de répondre Sans objet
9. Est-ce que les méthodes utilisées pour combiner les résultats des études sont appropriées? Pour ce qui est des résultats regroupés, un test devrait d'abord être effectué afin d'évaluer l'homogénéité des études pour ensuite déterminer si elles peuvent être combinées (p. ex. test du chi carré d'homogénéité, I ²). En présence d'hétérogénéité, le modèle des effets aléatoires devrait être utilisé ou s'il est pertinent sur le plan clinique, la combinaison des résultats devrait être prise en considération (c.-à-d. est-ce raisonnable de les combiner?)	Oui Non Impossible de répondre Sans objet
10. La possibilité d'un biais de publication a-t-elle été évaluée? L'évaluation du biais de publication devrait inclure une variété de représentations graphiques (p. ex. diagramme de dispersion, autres tests disponibles) ou de tests statistiques (p. ex., test de régression d' Egger).	Oui Non Impossible de répondre Sans objet
11. Les conflits d'intérêts ont-ils été divulgués? Les sources potentielles de soutien devraient être clairement énoncées dans la revue systématique et les études retenues.	Oui Non Impossible de répondre Sans objet

Une linguiste (JAB) a également participé à la production de la version préofficielle de la version franco-canadienne de l'outil AMSTAR, comme suggère la méthode Vallerand (1989).¹² Bien qu'une des cochercheuses expertes de l'étude (AS) possède une maîtrise en linguistique et qu'une orthophoniste (CD) formée en tant que

thérapeute du langage siégeait aux différents comités qui ont mené à la version préofficielle de l'outil AMSTAR, la présence d'une linguiste professionnelle (JAB) a considérablement complété l'expertise des divers comités d'experts.

Étape 5 : traduction inversée vers l'anglais de la version préofficielle

Afin de rendre la traduction franco-canadienne de l'outil AMSTAR officielle (Boîte 1), une étape additionnelle de traduction inversée vers l'anglais de la version préofficielle et effectuée par une cochercheuse (JK) bilingue, mais de langue maternelle anglophone en fin de processus. Cette étape supplémentaire a été inspirée par la méthodologie de Beaton et al. (2000).²¹

Études de cohérence interne, de fiabilité interjuges et de validation factorielle

La fiabilité interjuges de la version franco-canadienne officielle de l'outil AMSTAR a été examinée à l'aide de quatre juges indépendants (NC, JT, CP, ST) qui ont évalué chacune des 20 revues systématiques : 1) en français; 2) publiées après les années 2000; et 3) portant sur l'efficacité d'intervention physique (non chirurgicale ni pharmacologiques). Les recherches d'articles ont été effectuées en novembre 2015 et en février 2016 par l'entremise de divers moteurs de recherche, notamment Google Scholar, Chercher + de la bibliothèque virtuelle de l'Université d'Ottawa et Annales de réadaptation et de médecine physique. Les mots-clés utilisés lors de la recherche sont : « revue systématique », « français », « réadaptation », « santé », « physiothérapie ». L'ensemble des revues éligibles a été sélectionné.

Le coefficient de corrélation intraclasse (CCI) de type (2,1) a été calculé afin d'évaluer la fiabilité interjuges entre l'ensemble des éléments de l'outil,²² tout en tenant compte de la variance pour les fois où il a été complété, pour la revue systématique évaluée et la marge d'erreur. Le coefficient kappa de Cohen a été utilisé, par ailleurs, pour mesurer la concordance entre les réponses des quatre juges pour chacun des éléments²² de l'outil AMSTAR. La cohérence interne a été mesurée par le calcul du coefficient alpha de Cronbach qui évalue si chacun des éléments reproduits de façon répétée et constante la mesure d'un même construit²², c'est-à-dire la qualité méthodologique des revues systématiques.

La validité de construit proposée et évaluée à l'aide d'une analyse factorielle²² a été utilisée afin de mesurer le nombre de dimensions/concepts expliqué(e)s par la variance commune sur des facteurs distincts. La méthode orthogonale a été retenue. La méthode d'analyse factorielle orthogonale est une procédure statistique où la variance commune des données se loge sur des facteurs indépendants les uns des autres. Une valeur de communauté de plus de 0,3 est considérée comme étant significative. La version 13 du logiciel Stata/IC a été utilisée pour l'analyse des données.

RÉSULTATS

Description des participants

Les différentes étapes de traduction et de validation transculturelle ont permis de produire une version satis-

faisante de l'outil AMSTAR. La description des membres des comités et des répondants ayant participé aux différentes étapes du processus de traduction et de validation transculturelle se trouve au Tableau 1. On peut constater que la majorité des experts bilingues ayant participé aux différentes étapes de l'étude étaient des professionnels de la santé dont la langue maternelle était le français, à l'exception d'une cochercheuse dont la langue maternelle était l'anglais (JK). La présence de cette dernière a aidé à bien saisir la portée de certaines expressions anglaises, afin de permettre à l'équipe de recherche de bien évaluer la concordance des énoncés dans les deux langues.

La totalité des futurs professionnels en physiothérapie ($n = 18$: 100 %) qui ont été sollicités pour participer à l'évaluation de la deuxième version expérimentale a accepté l'invitation. L'échantillon était composé en majorité de femmes (88,9 %). La totalité des participants avait le français comme langue maternelle (100 %) et tous étaient également bilingues, puisqu'il s'agissait d'un critère d'admission du programme de physiothérapie. Tous ces futurs professionnels de la réadaptation avaient déjà évalué au moins une revue systématique (18/18) et étaient familiarisés avec l'utilisation de l'outil AMSTAR dans l'évaluation de la qualité méthodologique d'une revue systématique (18/18). Cette activité d'évaluation était obligatoire durant de leur formation universitaire.

Élaboration de la version expérimentale

Lors de la rencontre avec le premier comité d'experts, il a été convenu d'atteindre un consensus sur la traduction des versions préliminaires entre chacun des membres en présence de discordances (Figure 1).

Évaluation de la version expérimentale

Pour l'ensemble des énoncés évalués, les résultats moyens obtenus à l'échelle d'ambiguïté se situent entre 1,0 et 1,4 point sur une échelle de 7 points où le niveau 1 signifiait une plus grande clarté. Le Tableau 2 fait état de ces résultats qui démontrent un niveau de clarté élevé ou de non-ambiguïté, ainsi qu'une compréhension adéquate de chacun des énoncés. Ce critère est en concordance avec la méthode proposée par Vallerand (1989).¹²

La dernière traduction inversée vers l'anglais de la version préofficielle de l'outil AMSTAR franco-canadien était identique à la version de l'outil AMSTAR en version originale anglaise. Cette dernière étape confirme également que la méthodologie employée et que la traduction franco-canadienne de l'outil AMSTAR officielle produite sont valides.

Études de cohérence interne, de fiabilité interjuges et de validité de construit factorielle

Le coefficient alpha de Cronbach représente une valeur de 0,91 (IC à 95 % : 0,83, 0,99). En ce qui concerne la fiabilité interjuges ($n = 4$), le coefficient de corrélation intraclasse (CCI) représente 0,61 (IC à 95 % : 0,29, 0,97)

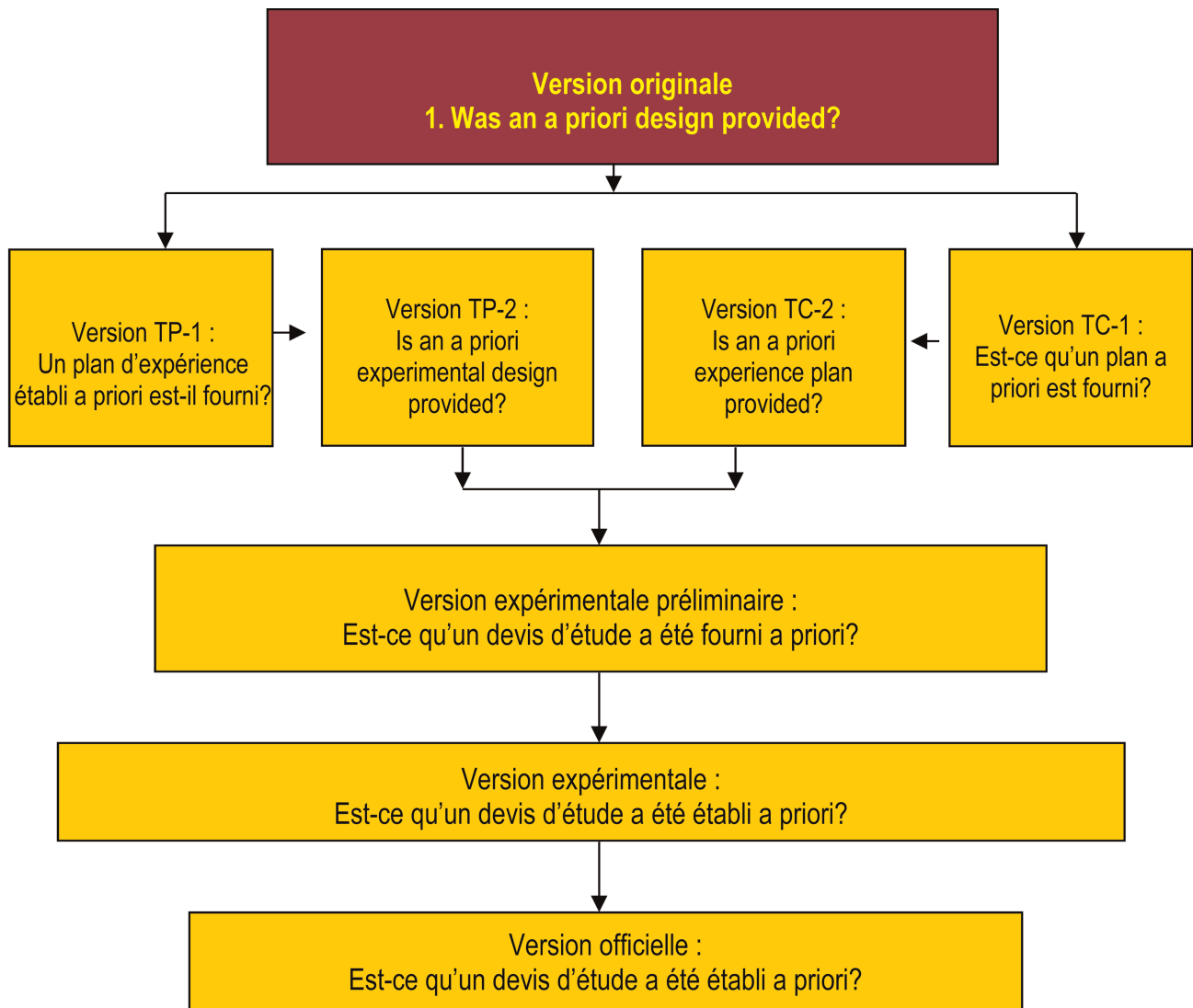


Figure 1 Processus de traduction selon la méthodologie proposée par Vallerand (1989)¹²

Tableau 2 Cote moyenne d'ambiguïté de chaque énoncé (score maximal /7)

Énoncé	Moyenne et σ
1	1,4 ± 1,0
2	1,2 ± 0,7
3	1,1 ± 0,4
4	1,4 ± 1,1
5	1,0 ± 0,1
6	1,1 ± 0,4
7	1,2 ± 0,6
8	1,2 ± 0,6
9	1,3 ± 0,9
10	1,1 ± 0,3
11	1,1 ± 0,5

pour le score total de l'outil AMSTAR. Cette fiabilité interjuges est considérée comme étant bonne compte tenu de la participation de quatre juges dans l'évaluation de la qualité méthodologique globale des 20 revues systématiques sélectionnées.

Selon l'échelle de McFayden, Webster et Maclaren (2006)²³ (Tableau 3), 82 % (9/11) des éléments ont montré, par leur coefficient kappa, une fiabilité interjuges entre les catégories excellente et moyenne (Tableau 4). En effet, les éléments 1, 5 et 7 ont été cotés à un niveau de fiabilité allant d'excellent à bon, selon les valeurs kappa correspondantes entre 0,71 et 1,0. Les éléments 2, 3, 4, 6, 8 et 10 ont été estimés à un niveau de fiabilité allant de bon à modéré, représentant des valeurs kappa entre 0,68 et 0,47. Seulement deux éléments (éléments 9 et 11) sont classés à un niveau de fiabilité pauvre avec des valeurs kappa respectives de 0,17 et 0,12.

Tableau 3 Échelle d'interprétation²³

Valeur kappa	Niveau de fidélité
<0,20	Pauvre
0,21–0,40	Acceptable
0,41–0,60	Modéré
0,61–0,80	Bon
0,81–1,00	Excellent

Tableau 4 Valeurs kappa pour les 20 SR

Élément	Kappa	Z	Prob > Z
1	1,00	10,95	0,000
2	0,6831	7,48	0,000
3	0,5628	6,17	0,00
4	0,5050	5,53	0,000
5	0,7151	7,83	0,000
6	0,4667	5,11	0,000
7	0,8377	9,18	0,000
8	0,7114	7,79	0,000
9	0,1667	1,83	0,0339
10	0,6190	6,78	0,000
11	0,1161	1,27	0,1017

L'analyse factorielle de l'outil AMSTAR a décelé qu'un seul facteur significatif qui représente le concept de la qualité méthodologique de RS. En effet, le tableau 5 révèle que la totalité des valeurs de communauté de l'analyse factorielle se loge sur un facteur unique et est plus grande que 0,3.

DISCUSSION

Cette étude visait à produire une traduction canadienne-française de l'outil AMSTAR et à examiner sa validité transculturelle, sa cohérence interne, sa fidélité interjuges et sa validité de construit factorielle. Les différentes étapes réalisées ont permis de produire une traduction valide.

Traduction et validité transculturelle

La traduction d'un instrument de mesure n'est pas un processus simple. Effectivement, plusieurs mots de la langue française partagent le même sens (p. ex. «aléatoire» et «au hasard»; «tenu secret», «ignoré», «inconnu» et «à l'insu de»; «assignation» et «répartition»; «d'admissibilité», «d'éligibilité» et «d'inclusion», etc.). Les participants des groupes d'experts ont dû faire des choix de mots selon leur habileté en français. Dans la présente étude, les participants (experts et chercheurs cliniciens) étaient tous bilingues, mais certains avaient une meilleure connaissance des anglicismes et des subtilités de la langue française. La présence d'une linguiste professionnelle et d'une orthophoniste expérimentée lors

Tableau 5 Analyse factorielle orthogonale

Élément	Facteur/dimension
1	X
2	0,4723
3	0,6404
4	0,7589
5	0,8126
6	0,3049
7	0,8102
8	0,9021
9	0,1084
10	0,3241
11	0,7020

X: L'analyse de variance n'a pas convergé pour l'élément 1 faute de variance dans les données.

Toutes les valeurs de communauté sont supérieures à 0,30.

de la rencontre des comités d'experts a permis de guider les choix d'expressions et de formulations, d'éviter des anglicismes et de nuancer les subtilités de la langue française. Cette force peut aussi présenter un défi. Étant donné la richesse de la langue française et la recherche du mot exact par les experts, il arrive fréquemment que des documents utilisant un niveau de français très formel ne soient pas faciles à comprendre par des francophones ayant des niveaux de langue variés et une origine culturelle diverse. Il semble que ce ne soit pas le cas ici. Les étudiants qui ont évalué la première version expérimentale de l'outil AMSTAR n'avaient pas tous une habileté équivalente en français, mais sont tous d'avis que les énoncés sont clairs. De plus, la dernière traduction inversée vers l'anglais de la version préofficielle de l'outil confirme également que la méthodologie employée et que la traduction franco-canadienne de l'outil AMSTAR officielle produite sont valides. Cette version franco-canadienne validée sera notamment utile aux professionnels de la santé issus de divers groupes culturels du Canada et partageant le français comme langue de travail.

Cohérence interne, fidélité interjuges et validité de construit factorielle

La version franco-canadienne de l'outil AMSTAR est unidimensionnelle. Elle offre une fiabilité interjuges modérée et une excellente cohérence interne. L'outil AMSTAR est une échelle méthodologique conçue pour évaluer la qualité méthodologique d'une revue systématique. Il se pourrait que le manque de clarté de certains énoncés traduits ne provienne pas seulement d'un problème linguistique, mais plutôt du niveau du jargon méthodologique spécifique au devis des essais cliniques, tel que 1) a priori, 2) le modèle des effets aléatoires, et 3) la littérature grise, etc. Le questionnaire démographique

aurait pu être amélioré en y ajoutant des questions liées à la clarté des énoncés de la version originale anglaise, ainsi qu'une question relative à la langue d'étude plutôt qu'à la langue maternelle.

La quatrième étape, qui mène à la version finale et qui fait appel aux utilisateurs potentiels, comporte plusieurs limites. La pénurie de professionnels de la santé dans certains établissements³ a contraint les chercheurs de l'étude actuelle à viser un échantillon de futurs professionnels de la santé afin d'augmenter le taux de réponse. Une représentation des étudiants dans toutes les disciplines de la réadaptation (p. ex. audiologie, ergothérapie, orthophonie) aurait été idéale, mais l'outil AMSTAR ne fait pas actuellement partie de toutes les formations à la pratique factuelle, donc la majorité de ces étudiant(e)s ne connaissent pas cet outil.

La version franco-canadienne de l'outil AMSTAR est un outil unidimensionnel qui offre une fiabilité inter-juges modérée et une excellente cohérence interne. La structure dimensionnelle de l'outil AMSTAR à l'aide de l'analyse factorielle effectuée dans la présente étude rejoint les conclusions d'études précédentes.^{5,9} La cohérence interne élevée obtenue confirme indirectement sa structure dimensionnelle. Cependant, la présente étude ne peut conclure que tous les 11 éléments constituant l'outil étaient nécessaires, puisque l'analyse factorielle n'a pas convergé pour l'élément 1 de l'outil AMSTAR.

La fiabilité inter-juges ($n = 4$) pour le score total de l'outil AMSTAR a été qualifiée comme modérée selon les valeurs du CCI et des coefficients kappa. Ces résultats sont cependant inférieurs à ceux de l'étude de Shea et collaborateurs (2009)⁵ et de Kang et collaborateurs (2012).⁸ Cependant, la plupart des éléments faibles de la version anglaise de l'outil AMSTAR⁵ correspondent aux mêmes éléments de la version franco-canadienne, à l'exception de l'élément 11. Ces différences peuvent être expliquées par le plus grand nombre de RS évaluées dans ces études, par des pairs de juges au lieu de quatre juges et par le plus grand choix nombre de répertoires établis de RS et de grande qualité (p. ex. «Cochrane Library»). De plus, l'élément 9, parmi les deux éléments les moins fiables, fait appel à des notions statistiques qui peuvent ne pas être parfaitement maîtrisées par certains juges. Pour sa part, l'élément 11 fait référence à l'identification des sources potentielles de conflit d'intérêts non seulement pour la revue systématique évaluée, mais également les études retenues dans cette revue. Le dernier critère peut s'avérer difficile à évaluer. Les résultats liés à la cohérence interne ne peuvent être comparés, puisqu'il n'existe aucune autre étude qui s'est intéressée à cette qualité méthodologique.

Idéalement, d'autres études de validation de l'outil AMSTAR devront être effectuées avec un plus grand échantillon de RS, mais leur nombre est limité au sein des écrits scientifiques publiés en langue française. Le

choix de sélection des RS publiées uniquement en français est justifié par l'utilisation de la terminologie française utilisée dans l'outil AMSTAR nouvellement traduit ainsi que par absence de barrière de langue.

CONCLUSION

Il est connu que les associations canadiennes professionnelles et les ordres professionnels provinciaux au Canada recommandent une pratique fondée sur les preuves scientifiques. Cependant, il existe peu d'outils de mesure validés dans la langue franco-canadienne pour mesurer la qualité des écrits scientifiques. Grâce aux cinq étapes de la méthode modifiée de Vallerand (1989),¹² la présente étude a permis de produire une version franco-canadienne de l'outil AMSTAR. La création de la version franco-canadienne de l'outil AMSTAR, dont la validité et la fiabilité sont acceptables, permettra de répondre au besoin grandissant de ce type d'outil compte tenu du nombre de professionnels de la santé francophones au Canada qui ne cesse d'augmenter. La francophonie sur le plan international pourrait éventuellement être également intéressée par cet outil.

MESSAGES CLÉS

Ce qui est déjà connu du sujet

L'outil AMSTAR est un outil pour évaluer la qualité méthodologique d'une revue systématique qui est validée dans la langue anglaise.

Ce que l'étude rajoute

Cette publication fournit une traduction rigoureuse franco-canadienne de l'outil AMSTAR.

RÉFÉRENCES

- Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Internet]. London: The Cochrane Collaboration; 2011. Disponible à : www.cochrane-handbook.org.
- Sharif MO, Janjua-Sharif FN, Ali H, et al. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. [Erratum in: Oral Health Dent Manag, 12]. [2]. Oral Health Dent Manag. 2013;12(1):9-16. Medline:23474576
- Brousseau L, Laroche C, Sutton A, et al. Une version franco-canadienne de la Physiotherapy Evidence Database (PEDro) Scale : L'Échelle PEDro. Physiother Can. 2015;67(3):232-9. <http://dx.doi.org/10.3138/ptc.2014-37F>. Medline:26839449
- Brouwers MC, Kho ME, Browman GP, et al, for the AGREE Next Steps Consortium. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. CMAJ. 2010;182(10):1045-52. <http://dx.doi.org/10.1503/cmaj.091714>. Medline:20513780
- Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20. <http://dx.doi.org/10.1016/j.jclinepi.2008.10.009>. Medline:19230606
- Gagnier JJ, Kellam PJ. Reporting and methodological quality of systematic reviews in the orthopaedic literature. J Bone Joint Surg Am. 2013;95(11):e771-7. <http://dx.doi.org/10.2106/JBJS.L.00597>. Medline:23780547
- Panic N, Leoncini E, de Belvis G, et al. Evaluation of the endorsement of the preferred reporting items for systematic reviews

- and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One*. 2013;8(12):e83138. <http://dx.doi.org/10.1371/journal.pone.0083138>. Medline:24386151
8. Kang D, Wu Y, Hu D, et al. Reliability and external validity of AMSTAR in assessing quality of TCM systematic reviews. *Evid Based Complement Alternat Med*. 2012;2012:732195. <http://dx.doi.org/10.1155/2012/732195>. Medline:22454679
 9. Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J*. 2010;4:84–91. Medline:21088686
 10. Suebnukarn S, Ngamboonsirisingh S, Rattanabanlang A. A systematic evaluation of the quality of meta-analyses in endodontics. *J Endod*. 2010;36(4):602–8. <http://dx.doi.org/10.1016/j.joen.2009.12.019>. Medline:20307731
 11. Hébert R, Bravo G, Voyer L. La traduction d'instruments de mesure pour la recherche gérontologique en langue française : critères méthodologiques et inventaires. *Rev Can Vieillesse*. 1994;13(3):392–405. <http://dx.doi.org/10.1017/S0714980800006206>.
 12. Vallerand RJ. Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques : implication pour la recherche en langue française. *Psychol Can*. 1989;30(4):662–80. <http://dx.doi.org/10.1037/h0079856>.
 13. Dussault M, Villeneuve P, Deaudelin C. L'échelle d'auto-efficacité des enseignants: validation canadienne-française du Teacher efficacy scale. *Rev Sci Edu*. 2001;27(1):181–94. <http://dx.doi.org/10.7202/000313ar>.
 14. Lysyk M, Brown GT, Rodrigues E, et al. Translation of the Leisure Satisfaction Scale into French: a validation study. *Occup Ther Int*. 2002;9(1):76–89. <http://dx.doi.org/10.1002/oti.157>. Medline:12375009
 15. René F, Casimiro L, Tremblay M, et al. Une version canadienne française du Lower Extremity Functional Scale (LEFS) : L'Échelle fonctionnelle des membres inférieurs (ÉFMI), partie I. *Physiother Can*. 2011;63(2):242–8. <http://dx.doi.org/10.3138/ptc.2010-11F>. Medline:22379265
 16. Pouchot J, Larbre JP, Lemelle I, et al, for the Pediatric Rheumatology International Trials Organisation Qualité de Vie en Rhumatologie Group. Validation of the French version of the Childhood Health Assessment Questionnaire (CHAQ) in juvenile idiopathic arthritis. *Joint Bone Spine*. 2002;69(5):468–81. [http://dx.doi.org/10.1016/S1297-319X\(02\)00432-3](http://dx.doi.org/10.1016/S1297-319X(02)00432-3). Medline:12477231
 17. Petit Larousse. *Le Petit Larousse illustré : Dictionnaire*. Paris : Édition Larousse; 2014.
 18. Petit Robert. *Le Petit Robert 1 : Dictionnaire alphabétique et analogique de la langue française*. Paris : Édition Le Robert; 2014.
 19. Fortin MF. *Fondements et étapes du processus de recherche : méthodes quantitatives et qualitatives*. 2e éd. Montréal: Chenelière Éducation; 2010.
 20. Termium Plus. Ottawa : Travaux Publics et Services Gouvernementaux Canada; c2014. Disponible à : <http://www.btb.termiumplus.gc.ca/>.
 21. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–91. <http://dx.doi.org/10.1097/00007632-200012150-00014>. Medline:11124735
 22. René F, Casimiro L, Tremblay M, et al. Fiabilité test retest et validité de construit de la version française de L'Échelle fonctionnelle des membres inférieurs (ÉFMI), partie II. *Physiother Can*. 2011;63(2):249–55. <http://dx.doi.org/10.3138/ptc.2010-12F>. Medline:22379266
 23. McFadyen AK, Webster VS, Maclaren WM. The test-retest reliability of a revised version of the Readiness for Interprofessional Learning Scale (RIPLS). *J Interprof Care*. 2006;20(6):633–9. <http://dx.doi.org/10.1080/13561820600991181>. Medline:17095441