

Traduction franco-canadienne de l'Assessment of Systematic Reviews Revised (AMSTAR 2) : validation transculturelle et fidélité interjuges

Heather Flowers, Ph. D.;* Paulette Guitard, Ph. D.;† Judy King, Ph. D.;‡ Elizabeth Fitzpatrick, Ph. D.;* Daniel Bérubé, Ph. D.;* Julie Alexandra Barette, M. Ling.;§ Dominique Cardinal, M.A.;¶ Sabrina Cavallo, Ph. D.;** Jennifer O'Neil, Ph. D.;†† Marylène Charette, M. Sc., pht.;‡‡ Laurence Côté, M. Sc., ergo.;†† Nalia Cecilia Gurgel-Juarez, DMD;†† Karine Toupin-April, Ph. D.;§§¶¶ Shirin M. Shallwani, M. Sc., pht.;†† Michelle Dorion, B. Sc., pht.;†† Prinon Rahman, M. Sc.;†† Maude Potvin-Gilbert, B. Sc.;†† Vanessa Bartolini, B. Sc.††

COMITÉ D'EXPERTS POUR LA RÉDACTION DE LA PREMIÈRE VERSION EXPÉRIMENTALE DE L'OUTIL AMSTAR 2

Krystina B. Lewis, Ph. D.;*** Rose Martini, Ph. D.;† Josée Lagacé, Ph. D.;* Roseline Galipeau, Ph. D.;†† Marie-Christine Ranger, M. Sc., ergo.;†† Fauve Duquette-Laplante, M. Sc. aud.;†† Marie-France Perrier, M. Sc. ortho.††

COMITÉ D'EXPERTS POUR L'ÉVALUATION DE LA 2^E VERSION EXPÉRIMENTALE DE L'OUTIL AMSTAR 2

Jacinthe Savard, Ph. D.;† Nicole Paquet, Ph. D.;‡ Jocelyne Tourigny, Ph. D.;*** Marie-Eve Bérubé, M. Sc., pht.;†† Hussein Ba Haroon, M. Sc., pht.;†† Patrick Duong, M. Sc., ergo.;†† Jacynthe Bigras, M. Sc. aud.;†† Julie Capistran, M. Sc., ergo.;†† Laurianne Loew, Ph.D.††

RÉSUMÉ

Objectif : produire une traduction en français canadien de l'AMSTAR 2, en affirmer la validité de contenu et en examiner la fidélité interjuges. **Méthodologie** : selon une approche adaptée de celle proposée par Vallerand, des traductions directes et renversées ont été effectuées. Ensuite, en examinant les traductions, un premier comité d'experts a créé la version expérimentale préliminaire. Celle-ci a été modifiée par un deuxième comité d'experts. Vingt futurs professionnels de la santé ont évalué la nouvelle version avec une échelle d'ambiguïté (de 1 à 7). Les cochercheurs principaux ont examiné les éléments problématiques pour affirmer la version préofficielle. Afin d'assurer la validité de contenu, une dernière rétrotraduction a été effectuée, validant une version officielle. Ensuite, 4 juges ont évalué 13 revues systématiques publiées en français à l'aide de la version officielle. Le coefficient kappa a été utilisé afin d'examiner la fidélité interjuges. **Résultats** : cette adaptation a permis la création d'une version franco-canadienne de l'AMSTAR 2. Son utilisation a démontré très peu d'ambiguïté (moyenne 1,15; ÉT 0,26) et une bonne fidélité interjuges (Kappa global > 0,64). **Conclusion** : la version franco-canadienne de l'AMSTAR 2 pourrait servir de soutien aux cliniciens, aux éducateurs et aux gestionnaires francophones au Canada lorsqu'ils cherchent à adopter une pratique factuelle.

Mots clés : outil d'évaluation, qualité méthodologique, revue systématique, traduction, validité transculturelle

*Programme d'audiologie et d'orthophonie, École des sciences de la réadaptation, Université d'Ottawa, Ottawa (Ontario) Canada; †Programme d'ergothérapie, École des sciences de la réadaptation, Université d'Ottawa, Ottawa (Ontario) Canada; ‡Programme de physiothérapie, École des sciences de la réadaptation, Université d'Ottawa, Ottawa (Ontario) Canada; §Département de linguistique, Université d'Ottawa, Ottawa (Ontario) Canada; ¶Consortium national de formation en santé (CNFS), Université d'Ottawa, Ottawa (Ontario) Canada; **Programme d'ergothérapie, École de réadaptation, Université de Montréal, Montréal (Québec) Canada; ††École des sciences de la réadaptation, Université d'Ottawa, Ottawa (Ontario) Canada; ‡‡École interdisciplinaire des sciences de la santé, Université d'Ottawa, Ottawa (Ontario) Canada; §§Institut de recherche du Centre hospitalier pour enfants de l'est de l'Ontario, Ottawa (Ontario) Canada; ¶¶Faculté de médecine et Faculté des sciences de la santé, Université d'Ottawa, Ottawa (Ontario) Canada; ***École des sciences infirmières, Université d'Ottawa, Ottawa (Ontario) Canada; †††Département des sciences infirmières, Université du Québec en Outaouais, Gatineau (Québec) Canada

Correspondance à : Heather Flowers, Faculté des sciences de la santé, École des sciences de la réadaptation, Université d'Ottawa, 451 Smyth, Pièce : RGN3071, Ottawa (Ontario) K1H 8M5 Canada; heather.flowers@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'intérêts à déclarer.

Remerciements : Les auteurs remercient le cabinet de la Vice-directrice à l'enseignement et 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 la D^{re} Lucie Brosseau, Bisma Rehman et Emmanuelle Harris-Ranger pour leur précieuse contribution.

ABSTRACT

Objective: Produce a French-Canadian translation of AMSTAR 2, affirm its content validity, and examine interrater reliability. **Methods:** Based on Vallerand's methodological approach, we conducted forward and parallel inverse-translations. Subsequently, an expert panel evaluated the translations to create a preliminary experimental French-Canadian version. A second expert panel examined this version and proposed additional modifications. Twenty future health professionals then rated the second experimental version for ambiguity on a scale (from 1 to 7). The principal co-investigators then reviewed the problematic elements and proposed a pre-official version. To ascertain content validity, a final back-translation was conducted resulting in the official version. Four judges evaluated 13 systematic reviews using the official French-Canadian version of AMSTAR 2. The Kappa coefficient was used to evaluate interrater reliability. **Results:** This rigorous adaptation enabled the development of a Franco-Canadian version of AMSTAR 2. Its application demonstrated low ambiguity (mean 1.15; SD 0.26) as well as good overall interrater reliability (total $\kappa > 0.64$) across all items. **Conclusion:** The French-Canadian version of AMSTAR 2 can now support francophone clinicians, educators, and managers in Canada as they undertake evidence-based practice.

Key Words: assessment tool; cross-cultural validity; methodological quality; systematic reviews; translation.

Les revues systématiques (RS) de bonne qualité devraient être accessibles à toute personne (étudiant en cursus professionnel, clinicien, chercheur ou gestionnaire) qui s'implique dans l'offre des services de santé nécessitant la synthèse des données probantes. Il est également souhaitable que ces derniers puissent évaluer et interpréter les synthèses des études scientifiques de façon rigoureuse afin de faire ressortir le plus haut niveau de preuves pour assurer une pratique factuelle, surtout à l'égard d'une approche diagnostique ou de l'efficacité d'une intervention^{1,2}. C'est d'autant plus important que les ordres professionnels canadiens dans le domaine de la santé recommandent une pratique fondée sur des données probantes. De telles exigences s'avèrent difficiles à satisfaire sans l'appui de ressources desservant l'avancement continu de la pratique et du savoir-faire.

Afin de vérifier la qualité méthodologique des écrits scientifiques², des chercheurs ont mis au point plusieurs outils. Un outil qui s'adresse à la fois aux cliniciens et aux experts est l'AMSTAR³, récemment amélioré et renommé AMSTAR 2^{4,5}. Quoique les qualités métrologiques de l'AMSTAR soient très enviables^{3,6-11}, certains auteurs ont souligné des faiblesses dans sa version initiale³, tel un manque de pertinence de certains éléments dans l'analyse de la qualité des recensions de synthèses^{5-7,9,12-19}.

L'AMSTAR 2 inclut l'ajout de cinq éléments ainsi que des précisions méthodologiques. Par exemple, le premier point de l'AMSTAR 2 met en valeur l'importance d'aligner la question de recherche et les critères d'inclusion avec les composantes PICO⁴. Il y a également des précisions (aux points 5 et 6) sur la rigueur méthodologique concernant la sélection d'articles et l'extraction des données par deux évaluateurs⁴. De nouveaux points (points 9, 12 et 13) portent sur l'évaluation du risque de biais, ce qui fournit plus d'information en vue d'une analyse critique et permet une meilleure interprétation de la qualité de l'ensemble des données probantes⁴. Contrairement à la version originale, l'AMSTAR 2 permet d'évaluer la qualité méthodologique des RS qui incluent des essais cliniques randomisés tout autant que de celles qui comprennent des essais cliniques non randomisés.

Des outils^{4,20} comme l'AMSTAR 2 devraient être disponibles dans la langue d'usage du clinicien pour s'assurer qu'un plus grand nombre d'entre eux puisse s'en servir dans leur pratique quotidienne^{12,21-41}. Une version franco-canadienne serait également utile pour les éducateurs œuvrant dans les programmes de santé francophones qui pourraient l'utiliser lors des formations professionnelles. Enfin, les gestionnaires de services francophones pourraient en promouvoir l'usage dans le développement de protocoles d'intervention et de lignes directrices dans plusieurs contextes en soins de santé.

Plusieurs études récentes ont examiné la validité convergente et la fidélité d'outils qui évaluent la qualité méthodologique des RS⁴¹⁻⁴⁴, dont l'AMSTAR, le R-AMSTAR⁴⁵ et l'AMSTAR 2, plus récemment. Une étude menée par Dosenovic⁴¹ et une autre par Pieper¹³ ont révélé des concordances entre les scores totaux de l'AMSTAR et ceux du R-AMSTAR, selon des corrélations Spearman de 0,88 ($p < 0,001$) et de 0,5 ($p = 0,14$) respectivement. L'étude ayant obtenu les plus fortes corrélations⁴¹ comprenait une évaluation de 97 RS par deux examinateurs tandis que la deuxième¹³ comptait 16 RS analysées par 5 examinateurs regroupés en 10 paires différentes¹³. Depuis, une étude menée par Lorenz a démontré une concordance entre l'AMSTAR et l'AMSTAR 2, selon une corrélation Spearman de 0,91 ($p < 0,001$)⁴⁴. Deux autres études^{43,44} comprenant 30 RS évaluées par 4⁴³ et 2⁴⁴ examinateurs, respectivement, ont révélé une corrélation Spearman de 0,84 ($p < 0,1$)^{43,44} entre l'AMSTAR 2 et le ROBIS⁴⁶.

En outre, les études menées par Dosenovic⁴¹ et Pieper¹³ ont démontré une meilleure fidélité interjuges dans l'évaluation des études systématiques avec l'AMSTAR qu'avec le R-AMSTAR. Dans l'étude de Dosenovic et coll.⁴¹, la moyenne des coefficients Kappa était de 0,43 (IC à 95 % : 0,36, 0,5) et de 0,51 (IC à 95 % : 0,33, 0,68) pour le R-AMSTAR et l'AMSTAR, respectivement. De même, l'étude de Pieper et coll.¹³ a révélé une plus faible fidélité selon les médianes des coefficients Kappa parmi les paires d'examineurs pour le R-AMSTAR ($\kappa = 0,45$; plage de 0,32 à 0,62) par rapport à l'AMSTAR ($\kappa = 0,52$; plage de 0,41 à 0,69). Tout récemment, Gates et coll.⁴² ont démontré une grande variation de fidélité interjuges de

l'AMSTAR 2 parmi trois groupes d'évaluateurs examinant les mêmes 30 RS dans divers centres au Canada, en Allemagne et au Portugal⁴². Les coefficients Kappa pour la confiance globale étaient de 0,42 (IC à 95 % : 0,12, 0,72), de 0,11 (IC à 95 % : -0,6, 0,29), et de 1,0 (IC à 95 % : parfait) selon le lieu. D'ailleurs, les études déjà mentionnées menées par Pieper⁴³ et Lorenz⁴⁴ ont révélé une fidélité médiocre ($\kappa = 0,30$; IC à 95 % : 0,17, 0,43) et modérée ($\kappa = 0,42$; IC à 95 % : 0,25, 0,59) pour la même analyse de confiance globale dans l'évaluation de 30 et 60 RS, respectivement.

Quoiqu'il existe une bonne concordance entre l'AMSTAR et l'AMSTAR 2, les taux de fidélité interjuges de l'AMSTAR 2 sont plus faibles que ceux de l'AMSTAR selon les études publiées. L'usage de l'AMSTAR 2 demeure tout de même pertinent compte tenu des améliorations importantes à l'évaluation de la qualité des études systématiques. Il est toutefois impératif d'entamer de nouvelles études avec l'AMSTAR 2 pour ajouter des preuves de sa fidélité interjuges. Actuellement, il existe une traduction franco-canadienne de l'outil AMSTAR^{3,11}. Puisque l'AMSTAR 2 est substantiellement différent de l'AMSTAR, une nouvelle traduction s'imposait. Le présent projet visait à produire une traduction de l'AMSTAR 2 en français canadien, d'en affirmer la validité de contenu et d'en examiner la fidélité interjuges.

MÉTHODOLOGIE

La présente étude a reçu une approbation déontologique de l'Université d'Ottawa, et elle a débuté à l'automne 2018. La méthode utilisée est présentée selon deux phases, soit : I) la production et la validation de contenu

(rétrotraduction) de la version franco-canadienne de l'AMSTAR 2 (figure 1); et II) l'évaluation de sa fidélité interjuges. Pour ce faire, les chercheurs ont dû recourir à deux comités d'experts et à des utilisateurs. Les critères de sélection étaient les suivants pour les comités d'experts et les utilisateurs, respectivement : les comités d'experts devaient inclure des chercheurs cliniciens qui avaient une bonne maîtrise de la langue française et qui résidaient au Canada. Les professeurs et les doctorants francophones de l'École des sciences de la réadaptation et de l'École des sciences infirmières ont reçu une invitation électronique générale. Les utilisateurs étaient des étudiants de physiothérapie (programme francophone) qui avaient complété leur cours de pratique factuelle dans lequel l'outil AMSTAR avait été enseigné. Ils ont tous reçu une lettre d'information et un formulaire de consentement à signer. Chaque participant a été compensé par une carte-cadeau de 5 \$ d'une compagnie de café. Afin de terminer la collecte de données, les membres des deux comités d'experts et les usagers consultés ont été invités à remplir un court questionnaire démographique.

Phase 1 : Traduction et validation de contenu

La traduction de l'AMSTAR 2⁴ a respecté une méthodologie quasi identique à celle d'une étude antérieure (complétée en 2017¹¹), d'après une version modifiée de la méthodologie conçue par Vallerand⁴⁷. La première phase contenait cinq étapes qui visaient la traduction transculturelle et la validation de contenu de l'AMSTAR 2. Les quatre premières étapes ont suivi de près la méthodologie de Vallerand⁴⁷ et visaient la création d'une version préofficielle. La cinquième étape représentait un

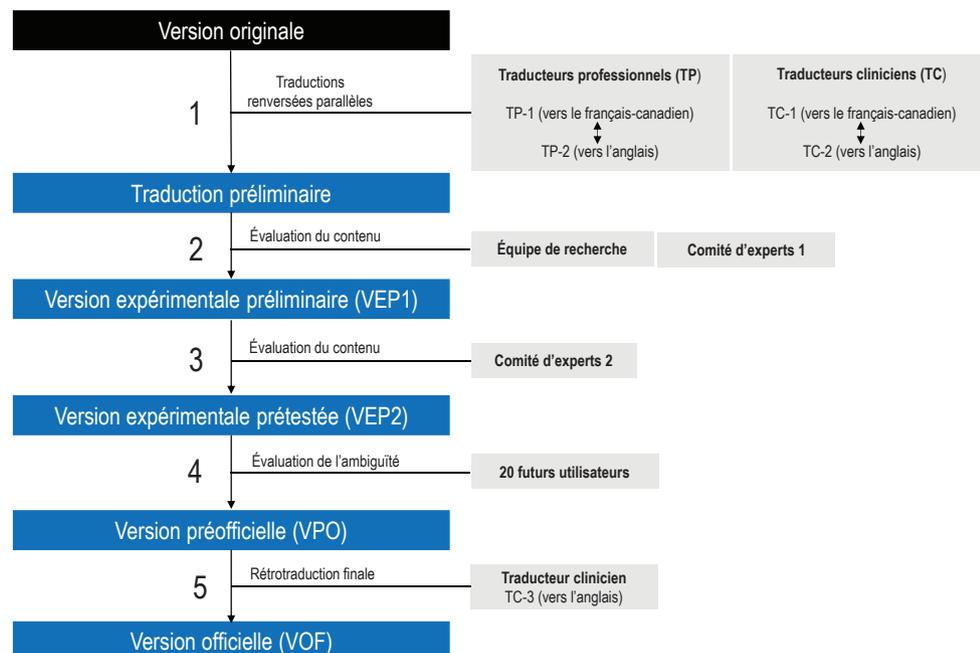


Figure 1 Processus de traduction selon Vallerand⁴⁷ (étapes 1 à 4) avec une rétrotraduction finale (étape 5)

ajout pour affirmer la validité de contenu grâce à un processus de rétrotraduction finale (version préofficielle traduite vers l'anglais pour valider le contenu culturel et conceptuel⁴⁸). Ainsi, les cinq étapes (voir la figure 1) sont :

- 1) la traduction renversée parallèle (TRP), soit la traduction de l'anglais au français (canadien) par une traductrice professionnelle bilingue (TP-1) et par une traductrice clinicienne et chercheuse bilingue (TC-1). Ces premières traductions étaient suivies de traductions renversées (vers l'anglais) faites par une deuxième traductrice professionnelle bilingue (TP-2) et une deuxième clinicienne et chercheuse bilingue (TC-2);
- 2) la rédaction de la version expérimentale préliminaire (VEP1) effectuée en comparant les cinq versions préliminaires de l'outil AMSTAR 2 (version originale, TP-1, TP-2, TC-1 et TC-2) par l'équipe de recherche et par le premier comité d'experts;
- 3) l'évaluation expérimentale prétestée (VEP2) de la VEP1 par le second comité d'experts qui devait confirmer ou modifier les énoncés selon un processus de consensus pour assurer l'interprétabilité selon les subtilités du français canadien;
- 4) l'évaluation de la version franco-canadienne prétestée (VEP2) selon une échelle d'ambiguïté allant de 1 à 7 (1 : très clair; 7 : très ambigu) par de futurs utilisateurs bilingues pour produire une version préofficielle (VPO);
- 5) la rétrotraduction vers l'anglais de la version préofficielle (VPO) par une chercheuse-clinicienne bilingue dont la langue maternelle est l'anglais, afin de produire la version officielle franco-canadienne (VOF) de l'AMSTAR 2.

Pour obtenir une description plus détaillée de chacune des étapes, le lecteur est invité à consulter l'étude antérieure ayant fait appel à la même méthodologie¹¹.

Phase 2 : Étude préliminaire de fidélité interjuges

La première étape consistait de repérer des RS dans les moteurs de recherche Érudit et Pubmed à l'aide des mots-clés « revue systématique », « français », « réadaptation », « santé », « physiothérapie ». Les critères d'inclusion se résumaient aux RS en français révisées par un comité de lecture et dont le texte était entièrement accessible en ligne.

Quinze RS publiées en français (dont deux incluant des méta-analyses) ont été sélectionnées et ensuite évaluées par des juges indépendants à l'aide de la version franco-canadienne officielle de l'AMSTAR 2 (liste bibliographique à l'annexe 1). Un total de quatre juges bilingues réparti en deux groupes a procédé à ces évaluations : i) deux juges

nouvellement formés (novices) et ii) deux juges expérimentés. Les deux juges novices ont évalué les 13 RS qui ne contenaient pas de méta-analyses et les deux juges expérimentés, les 15 RS répertoriées.

Trois autres RS ont été sélectionnées pour former les juges novices à l'utilisation de l'outil AMSTAR 2. Puisque le bassin de RS rédigées en français était très restreint, les trois RS faisant partie de la formation avaient été rédigées en langue anglaise. Les juges novices ont utilisé la version franco-canadienne, mais ont également consulté l'outil AMSTAR 2 en langue anglaise pour évaluer individuellement chaque RS, puis ont comparé leurs cotations et sont arrivés à un consensus pour chaque élément. Lors de cet exercice, si un consensus n'était pas possible entre les deux juges novices, la décision était tranchée par une chercheuse expérimentée après une discussion et une tentative de résolution.

Les 16 éléments de l'AMSTAR 2 sont cotés par « non », « oui », « oui en partie » ou « aucune méta-analyse effectuée » (non applicable). Les quatre juges ont coté les éléments selon une valeur numérique polytomique (c'est-à-dire 0 = non; 1 = oui ou « oui en partie »; 2 = « aucune méta-analyse effectuée »). La cotation « oui en partie » (qui figure uniquement dans les éléments 2, 4, 7, 8 et 9) n'a pas été retenue dans cette étude, afin de simplifier la tâche des juges novices et d'uniformiser le système de cotation pour l'ensemble des éléments.

Le pourcentage d'accord et le coefficient Kappa (Kappa de Fleiss pour les comparaisons entre plus de deux juges^{49,50}) ont été analysés pour chaque élément de l'AMSTAR 2 afin de déterminer la concordance entre les réponses des quatre juges (pour les 13 RS) et des deux juges expérimentés (pour les 15 RS). La fidélité interjuges de l'AMSTAR 2 a été qualifiée en fonction de l'interprétation proposée par McFayden, Webster et Maclaren⁵¹ (< 0,20 : *pauvre*; 0,21–0,40 : *médiocre*; 0,41–0,60 : *modérée*, 0,61–0,80 : *bonne*; 0,81–1,00 : *excellente*). Les logiciels SPSS⁵² et SAS 9.4 (SAS Institute, Cary NC) ont été utilisés pour l'analyse des données.

RÉSULTATS

Phase 1 : Description des participants

Seize experts et 20 utilisateurs ont participé aux différentes étapes du processus de traduction transculturelle (tableau 1). Tous les experts étaient des professionnels bilingues parmi lesquels 13 considéraient le français comme leur langue maternelle. Parmi les utilisateurs, 18 le considéraient comme leur langue maternelle.

Phase 1 : Traduction et élaboration des versions préliminaires (étapes 1, 2 et 3)

À la suite des versions produites par les traducteurs professionnels et cliniciens chercheurs (étape 1 : TRP),

Tableau 1 Données sociodémographiques des participants au processus de traduction

Caractéristiques	Comité d'experts 1 (n = 7)	Comité d'experts 2 (n = 9)	Futurs utilisateurs (n = 20)
Âge, moyenne (écart-type) ans	43 (11,87)	43 (14,59)	24,85 (2,37)
Sexe	Femmes : 7 Hommes : 0	Femmes : 7 Hommes : 2	Femmes : 13 Hommes : 6 Donnée manquante : 1
Langue maternelle, n	Français : 6 Italien : 1	Français : 7 Anglais : 1 Arabe : 1	Français : 18 Anglais : 1 Français et anglais : 1
Profession	Orthophoniste : 1 Audiologistes : 2 Ergothérapeutes : 2 Infirmières : 2	Physiothérapeutes : 4 Audiologiste : 1 Ergothérapeutes : 3 Infirmière : 1	Étudiants en physiothérapie : 20
Utilisé AMSTAR (1 ou 2)	Oui : 2 Non : 5	Oui : 4 Non : 5	Oui : 17 Non : 2 Donnée manquante : 1
Langue de travail	Français : 5 Anglais : 1 Français et anglais : 1	Français : 6 Anglais : 3	Français : 18 Anglais : 1 Français et anglais : 1
Années d'études post-collégiales, moyenne (écart-type)	9,71 (2,36)	11,67 (5,47)	6,35 (1,39)
Expérience en consultation de revue systématique, n	Oui : 7	Oui : 9	Oui : 17 Donnée manquante : 3

Version originale	Did the review authors perform study selection in duplicate?	
Étape 1	TP-1: Est-ce que les auteurs de la revue effectuent la sélection des études en double?	TC-1: Est-ce que les auteurs de la revue ont écarté les études dédoublées?
	TP-2: Do the authors of the review complete a selection of duplicate studies?	TC-2: Have authors removed study duplicates?
Étape 2	Est-ce que les auteurs de la revue ont effectué la double sélection des études?	
Étape 3	Est-ce que la sélection des études de la revue systématique a été effectuée par au moins deux évaluateurs?	
Étape 4	Aucune modification proposée	
Étape 5	Aucune modification proposée	

Figure 2 Exemple de modifications dans le processus de traduction pour le point 5 de l'AMSTAR 2

il y a eu des rencontres de chacun des comités d'experts d'une durée de trois heures, afin d'adopter des précisions et des simplifications (voir des exemples de traductions divergentes à l'annexe 2). Ces comités d'experts ont

contribué respectivement à l'élaboration de deux versions expérimentales (étape 2 : VEP1 et étape 3 : VEP2; voir un exemple de modifications proposées tout au long du processus à la [figure 2](#)).

Phase 1 : Élaboration de la version préofficielle (étape 4)

Les futurs utilisateurs ayant participé à cette étape étaient tous bilingues (critère d'admission du programme de maîtrise professionnelle de physiothérapie de l'Université d'Ottawa) et étaient en majorité des femmes (13; 65 %). La majorité considérait le français canadien comme leur langue maternelle (18; 90 %) (tableau 1). Pour l'ensemble des énoncés évalués, les résultats obtenus à l'échelle d'ambiguïté montrent un niveau de clarté très élevé des 16 éléments de la VEP2 (moyenne totale de $1,15 \pm 0,26$) (tableau 2), ainsi qu'une compréhension adéquate de tous les énoncés (annexe 3).

Phase 1 : Validation de contenu de la version officielle (étape 5)

La rétrotraduction de la version préofficielle de l'outil franco-canadien vers l'anglais a permis d'obtenir une version quasi identique à la version originale anglaise. Les divergences correspondaient surtout à des synonymes de certains termes. Cependant, l'expression « des études d'intervention non randomisées (EINR) » a causé des problèmes de traduction vers l'anglais de « non-randomized controlled study » au lieu de « non-randomized study of interventions ». Outre cette particularité, cette dernière étape a confirmé que la méthodologie employée ainsi que le contenu de la traduction produite étaient valides (VOF à l'annexe 4).

Phase 2 : Étude préliminaire de fidélité inter-juges

À la suite de l'obtention de la VOF de l'AMSTAR 2 en français canadien, l'étude de fidélité inter-juges a été entamée à partir des 15 RS sélectionnées. Les résultats de l'étude de fidélité comprenaient les évaluations faites par les juges expérimentés (15 RS) et celles faites par

l'ensemble des quatre juges (novices et expérimentés) (13 RS). Pour les juges expérimentés (tableau 3), la fidélité inter-juges selon les coefficients Kappa était d'excellente à modérée pour huit des 12 éléments analysés et de médiocre à pauvre pour quatre éléments. Cependant, 10 des 16 éléments de l'AMSTAR 2 ont révélé des pourcentages d'accord plus élevés que 80 %. Pour l'ensemble des juges (tableau 4), les pourcentages d'accord étaient moins

Tableau 3 Valeurs Kappa pour chaque élément de l'outil AMSTAR 2 évalué par deux juges expérimentés ($n = 15$ articles)

Élément	Kappa	Valeur p	Taux d'accord, n° (%)
1	0,01	0,93	6 (40)
2*	Indéterminé	Indéterminé	14 (93)
3	0,76	0,01	14 (93)
4*	Indéterminé	Indéterminé	14 (93)
5	0,28	0,11	10 (67)
6	0,45	0,06	11 (73)
7*	Indéterminé	Indéterminé	14 (93)
8	0,57	0,02	12 (80)
9	0,42	0,10	13 (87)
10*	Indéterminé	Indéterminé	12 (80)
11	0,80	0,01	14 (93)
12	1,00	0,01	15 (100)
13	0,18	0,21	8 (43)
14	0,13	0,26	9 (60)
15	1,00	0,01	15 (100)
16	0,81	0,01	14 (93)
Total	0,64	0,01	80,83 %

* Scores constants pour 1 juge (même cotation à chaque RS évaluée)

Tableau 4 Valeurs Kappa de Fleiss pour chaque élément de l'outil AMSTAR 2 évalué par quatre juges ($n = 13$ articles)

Éléments	Kappa	Valeur p	Taux d'accord, n° (%)
1*	Indéterminé	Indéterminé	3 (23)
2*	Indéterminé	Indéterminé	10 (77)
3	0,40	0,01	7 (54)
4†	Indéterminé	Indéterminé	12 (92)
5	0,40	0,01	6 (46)
6	0,53	0,01	7 (54)
7†	Indéterminé	Indéterminé	11 (85)
8	0,37	0,01	6 (46)
9	0,62	0,01	11 (85)
10†	Indéterminé	Indéterminé	11 (85)
11	0,50	0,01	10 (77)
12	0,51	0,01	10 (77)
13	0,14	0,20	3 (23)
14	0,21	0,05	6 (46)
15	0,50	0,01	10 (77)
16	0,78	0,01	11 (85)
Total	0,65	0,01	64,42 %

* Scores constants pour un juge

† Scores constants pour deux juges

Tableau 2 Score moyen d'ambiguïté des 16 éléments principaux de l'outil AMSTAR 2 (échelle d'ambiguïté de 7 points, de 1-très clair à 7- très ambigu)

Éléments AMSTAR 2	Moyenne	ÉT
1	1,21	0,71
2	1,95	1,13
3	1,42	0,77
4	1,40	0,82
5	1,00	0,00
6	1,16	0,37
7	1,00	0,00
8	1,00	0,00
9	1,06	0,24
10	1,00	0,00
11	1,00	0,00
12	1,00	0,00
13	1,00	0,00
14	1,18	0,53
15	1,00	0,00
16	1,00	0,00
Total	1,15	0,26

élevés que ceux des juges expérimentés séparément. Néanmoins, parmi les 11 analyses Kappa, 6 éléments ont atteint un niveau de fidélité de *modéré à excellent*.

DISCUSSION

Cette étude visait à produire une traduction franco-canadienne de l'AMSTAR 2 et à en affirmer la validité de contenu ainsi que la fidélité interjuges. L'outil AMSTAR 2 a été traduit selon un processus rigoureux, et la version franco-canadienne démontre très peu d'ambiguïté entre utilisateurs. De plus, la rétrotraduction finale vers l'anglais a révélé une absence de problèmes conceptuels lors de sa comparaison avec la version originale, ce qui suggère une bonne validité de contenu⁴⁸. La fidélité interjuges s'est avérée modérée pour la plupart des éléments parmi les juges novices et experts. Les chercheurs considèrent donc que la version franco-canadienne offre de bonnes propriétés métrologiques et est prête à l'emploi pour la formation des professionnels en soins de santé et en milieu clinique.

Lors des discussions avec les groupes d'experts, leurs habiletés en français et leurs connaissances méthodologiques ont influencé le choix des mots à retenir pour la version finale. Bien que les experts soient tous bilingues, certains avaient une meilleure connaissance des emprunts linguistiques, des anglicismes et des subtilités des dialectes canadiens de langue française. Le vocabulaire riche de la langue française et celui propre au domaine des RS ont par ailleurs mené à certains débats lors des réunions d'experts au sujet de synonymes, d'expressions spécifiques ou de tournures de phrase. La participation d'une linguiste et d'orthophonistes expérimentées a permis de renforcer la sélection d'expressions et de formulations compatibles avec l'usage du français au Canada dans les milieux de la santé.

L'évaluation de la deuxième version expérimentale (VEP2) de l'AMSTAR 2 par les futurs utilisateurs (étudiants en physiothérapie) a démontré que les éléments étaient très clairs, compte tenu des faibles scores obtenus à l'échelle d'ambiguïté. Tout au long du processus de traduction, les membres des différents groupes d'experts ont respecté l'esprit de la version anglaise de l'AMSTAR 2 et n'ont pas tenté de l'améliorer. Par exemple, le mot *participants* aurait été préféré à l'expression *patients* dans certains contextes. Par contre, certains utilisateurs ont jugé que le vocabulaire utilisé dans la traduction était de niveau trop élevé. Les expressions *littérature grise*, *mise à l'insu* et *compte rendu* leur étaient moins familières. Il faut noter que les participants qui évaluaient la VEP2 étaient des étudiants à la maîtrise professionnelle, alors que l'AMSTAR 2 est plutôt destiné à des chercheurs. De plus, la majorité d'entre eux étaient plus familiers avec la première version de l'outil (soit l'AMSTAR) et n'avaient qu'une

utilisation limitée des outils visant l'évaluation des qualités méthodologiques de la recherche. Même les cliniciens expérimentés ne sont souvent pas habitués à faire ce genre d'évaluation. Cet outil pourrait les aider à évaluer des RS pour mieux en comprendre les résultats et les adopter dans leur pratique. Il est donc important, lors de la formation de futurs cliniciens, de les familiariser avec ce genre d'outils pour qu'ils se sentent mieux outillés pour les utiliser et soutenir une pratique reposant sur des données probantes. Ainsi, une étude comme celle-ci permet de former les futurs cliniciens à l'évaluation des données probantes pendant leur cursus professionnel francophone.

L'évaluation de la fidélité interjuges de la version franco-canadienne de l'AMSTAR 2 était pertinente afin d'en démontrer la fiabilité et d'en favoriser l'adoption et l'utilisation auprès des professionnels francophones en soins de santé. La fidélité interjuges pour l'ensemble des éléments de la version franco-canadienne de l'AMSTAR 2 est considérée comme relativement bonne selon les valeurs Kappa obtenues des scores totaux. Lorsque les éléments individuels sont examinés, plus des deux tiers et de la moitié ont montré une fidélité interjuges d'*excellente à modérée* pour les juges expérimentés et pour les quatre juges (comprenant les résultats des juges expérimentés), respectivement. Ces résultats sont d'autant plus souhaitables que l'emploi des statistiques Kappa est très conservateur.

Par rapport à plusieurs études de fidélité interjuges en langue anglaise⁴²⁻⁴⁴, les résultats de cette étude sont convaincants, surtout à l'égard des deux juges experts. Néanmoins, parmi les RS des trois études de langue anglaise, il y avait une haute fréquence d'essais cliniques randomisés (de 50⁴³, de 70⁴⁴ et de 100⁴² pour cent). Il se peut que les 15 RS choisies soient plus homogènes et d'une qualité inférieure aux RS dans les études de langue anglaise. Par exemple, il a été difficile de trouver des RS contenant des méta-analyses. De plus, les deux juges experts avaient de l'expérience dans la réalisation d'études systématiques avec et sans méta-analyses. Le fait d'avoir des RS de qualité inférieure a pu faciliter le choix des cotes surtout pour les deux juges experts ce qui a donc donné un taux d'accord plus élevé.

En outre, les études de fidélité réalisées en anglais ont révélé une grande variabilité de certaines cotations selon les valeurs Kappa parmi les 5 groupes d'examineurs (dont 3 dans l'étude internationale de Gates et coll.⁴²), notamment pour les points suivants : 1 (κ variant de 0,08⁴² à 0,87⁴²), 3 (-0,06⁴² à 0,65^{42,43}), 5 (0,33⁴³ à 1,00⁴²), 8 (-0,09⁴² à 0,63⁴²) et 13 (0,10⁴² à 0,87⁴²). En particulier, l'étude de Gates et coll.⁴² contenait trois groupes d'examineurs ayant des niveaux d'expérience en lecture et interprétation de RS variés. De plus, les examineurs des trois études de langue anglaise ont dû utiliser plusieurs outils en parallèle (l'AMSTAR 2 et le ROBIS^{43,44}; l'AMSTAR,

l'AMSTAR 2 et le ROBIS⁴²⁾ lors de leurs évaluations des RS, ce qui aurait pu compliquer l'application des critères de l'AMSTAR 2 de manière constante. La nouvelle échelle de l'AMSTAR 2 franco-canadienne possède des qualités métrologiques comparables à celles signalées dans la version originale de l'outil AMSTAR 2. En effet, tout comme dans la présente étude, il existe une très grande variabilité des valeurs Kappa dans les différents éléments de l'AMSTAR 2^{43,44}. Des démarches méthodologiques et un engagement pour la formation de la réflexion critique des futurs professionnels devraient atténuer l'usage variable des outils comme l'AMSTAR 2 et promouvoir la fidélité écologique et la capacité des cliniciens francophones du Canada à interpréter correctement les études pour qu'ils puissent exercer une pratique fondée sur des données probantes.

LIMITES

Parmi les limites de la présente étude, la paire de juges novices ne maîtrisait peut-être pas à 100 % la terminologie liée aux RS. En général, on suggère des évaluateurs qui sont experts à la fois du sujet, de la méthodologie et des statistiques caractéristiques des RS¹⁴. Toutefois, l'analyse de l'utilisation par des juges novices a fourni des renseignements importants pour la formation continue de nouveaux cliniciens chercheurs. Un plus grand nombre de juges et de revues rédigées en français canadien (et non en français européen) aurait également été souhaitable. Autrement dit, un échantillon comprenant plus de juges de groupes variés (étudiants en formation dans plusieurs programmes de réadaptation et chercheurs cliniciens) aurait offert une vue plus écologique de la performance des examinateurs dans la cotation des points de l'AMSTAR 2.

Orientations futures

À l'avenir, il sera pertinent d'inclure des RS provenant d'autres disciplines en sciences de la santé, en dehors de la physiothérapie. C'est à dire que d'autres études de fidélité de l'AMSTAR 2 devront être effectuées avec un plus grand échantillon de RS en langue française provenant d'une sphère plus large de soins en santé. De telles améliorations permettront de mieux saisir la portée et la pertinence de l'utilisation fiable de la version franco-canadienne de l'AMSTAR 2.

CONCLUSION

Les associations et les ordres professionnels canadiens du domaine de la santé recommandent une pratique fondée sur des preuves scientifiques. Il s'agit d'un objectif difficile à atteindre sans le développement continu des ressources et du savoir-faire. À cet effet, l'analyse critique de la qualité méthodologique des RS demeure une habileté primordiale à acquérir chez les futurs professionnels de la santé, les éducateurs, les cliniciens

et les gestionnaires. Les adaptations transculturelles de nouveaux outils et l'amélioration d'outils éprouvés comme l'AMSTAR assureront la mise en œuvre d'une pratique factuelle à la hauteur des connaissances. La présente étude a permis la production d'une version franco-canadienne rigoureuse de l'outil AMSTAR 2. Sa validité de contenu ainsi que sa fidélité parmi des juges novices et experts ont été démontrées. La version franco-canadienne de l'AMSTAR 2 est prête à être utilisée dans la formation des professionnels en soins de santé et en milieu clinique et pour ainsi soutenir l'interprétation et l'usage des données probantes dans la pratique.

MESSAGES CLÉS

Ce qui est déjà connu du sujet

L'AMSTAR 2 en français canadien est un outil adéquat pour évaluer la qualité méthodologique des RS, et il est fidèle à la version en langue anglaise.

Ce que l'étude ajoute

La présente étude présente une version franco-canadienne de l'outil AMSTAR 2 après une traduction rigoureuse. L'AMSTAR 2 franco-canadien possède une fidélité interjuges comparable à celle de la version originale.

RÉFÉRENCES

- Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions 5.1.0 [Internet]. London : The Cochrane Collaboration; 2011. Disponible à : <https://handbook-5-1.cochrane.org/>.
- Sharif MO, Janjua-Sharif FN, Ali H et coll. Systematic reviews explained: AMSTAR-how to tell the good from the bad and the ugly. [Erratum in: Oral Health Dent Manag, 12]. Oral Health Dent Manag. 2013;12(1):9–16. Medline:23474576
- Shea BJ, Grimshaw JM, Wells GA et coll. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC. 2007;7(10):1–7. <https://doi.org/10.1186/1471-2288-7-10>. Medline:17302989
- Shea BJ, Wells GA, Thuku M et coll. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358–66. <https://doi.org/10.1136/bmj.j4008>. Medline:28935701
- Wu X, Sun H, Zhou X et coll. Quality assessment of systematic reviews on total hip or knee arthroplasty using mod-AMSTAR. BMC Med Res Methodol. 2018;18(1):30. <https://doi.org/10.1186/s12874-018-0488-8>. Medline:29548276
- Dosenovic S, Kadic AJ, Vucic K et coll. Comparison of methodological quality rating of systematic reviews on neuropathic pain using AMSTAR and R-AMSTAR. BMC Med Res Methodol. 2018;18(1):37. Medline:29739339
- Panic N, Leoncini E, de Belvis G et coll. 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://doi.org/10.1371/journal.pone.0083138>. Medline:24386151
- Kang D, Wu Y, Hu D et coll. Reliability and external validity of AMSTAR in assessing quality of TCM systematic reviews. Evid Based Complement Alternat Med. 2012;2012:732195. <http://doi.org/10.1155/2012/732195>. Medline:22454679
- Kung J, Chiappelli F, Cajulis OO et coll. 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. <https://doi.org/10.2174/1874210601004020084>. Medline:21088686
10. Shea BJ, Hamel C, Wells GA et coll. 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://doi.org/10.1016/j.jclinepi.2008.10.009>. Medline:19230606
 11. Brosseau L, Laroche C, Guitard P et coll. La version franco-canadienne de l'outil Assessment of Multiple Systematic Reviews (AMSTAR). *Physiother Can*. 2017;69(1):20–9. <http://doi.org/10.3138/ptc.2015-80F> Medline:28154441
 12. Shetty K, Raaen L, Khodyakov D et coll. Evaluation of the work loss data institute's official disability guidelines. *J Occup Environ Med*. 2018;60(3):e146–51. <http://doi.org/10.1097/JOM.0000000000001230>. Medline:29135837
 13. Pieper D, Jacobs A, Weikert B et coll. Inter-rater reliability of AMSTAR is dependent on the pair of reviewers. *BMC Med Res Methodol*. 2017;17(1):1–8. <https://doi.org/10.1186/s12874-017-0380-y>. Medline:28693497
 14. Pieper D, Koengen N, Breuing J et coll. How is AMSTAR applied by authors – a call for better reporting. *BMC Med Res Methodol*. 2018;18(1):56–63. <https://doi.org/10.1186/s12874-018-0520-z>. Medline:29914386
 15. Thompson C, Schabrun S, Romero R et coll. Contributing to chronic ankle instability: a systematic review and meta-analysis of systematic reviews. *Sports Med*. 2017;48(1):1–17. <https://doi.org/10.1007/s40279-017-0781-4>. Medline:28887759
 16. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. *Syst Rev*. 2016;5(1):58. <https://doi.org/10.1186/s13643-016-0237-1>. Medline:27072548
 17. Wegewitz U, Weikert B, Fishta A et coll. Resuming the discussion of AMSTAR: what can (should) be made better? *BMC Med Res Methodol*. 2016;16(1):111–19. <https://doi.org/10.1186/s12874-016-0183-6>. Medline:27566440
 18. Faggion CM. Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor. *BMC Med Res Methodol*. 2015;15(1):63. <https://doi.org/10.1186/s12874-015-0062-6>. Medline:26268372
 19. Popovich I, Windsor B, Jordan V et coll. Methodological quality of systematic reviews in subfertility: a comparison of two different approaches. *PLoS One*. 2012;7(12):e50403. <https://doi.org/10.1371/journal.pone.0050403>. Medline:23300526
 20. Pollock M, Fernandes RM, Hartling L. Evaluation of AMSTAR to assess the methodological quality of systematic reviews in overviews of reviews of healthcare interventions. *BMC Med Res Methodol*. 2017;17(1):48. <https://doi.org/10.1186/s12874-017-0325-5>. Medline:28335734
 21. Militello LK, Gance-Cleveland B, Aldrich H et coll. A methodological quality synthesis of systematic reviews on computer-mediated continuing education for healthcare providers. *Worldviews on Evid Based Nurs*. 2014;11(3):177–86. <https://doi.org/10.1111/wvn.12041>. Medline:24865984
 22. Wasiak J, Tyack Z, Ware R et coll. Poor methodological quality and reporting standards of systematic reviews in burn care management. *Int Wound J*. 2017;14:754–63. <https://doi.org/10.1111/iwj.12692>. Medline:27990772
 23. Pundir J, Charles D, Sabatini L et coll. Overview of systematic reviews of non-pharmacological interventions in women with polycystic ovary syndrome. *Hum Reprod Update*. 2019;25(2):243–56. <https://doi.org/10.1093/humupd/dmy045>. Medline:30608609
 24. Park M, Thi-Thanh-Tinh G, Lee M et coll. Patient- and family-centered care interventions for improving the quality of health care: a review of systematic reviews. *Int J Nurs Stud*. 2018;87:69–83. <https://doi.org/10.1016/j.nurstu.2018.07.006>. Medline:30056169
 25. Zhang F, Shen A, Jin Y et coll. The management strategies of cancer-associated anorexia: a critical appraisal of systematic reviews. *BMC Complement Altern Med*. 2018;18(1):1–9. <https://doi.org/10.1186/s12906-018-2304-8>. Medline:30092794
 26. Kowalczyk M, Adamich J, Simunovic N et coll. Methodological quality of systematic reviews addressing femoroacetabular impingement. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(9):2583–9. <https://doi.org/10.1007/s00167-014-3151-x>. Medline:25037984
 27. Thompson C, Schabrun S, Romero R et coll. Factors contributing to chronic ankle instability: a systematic review and meta-analysis of systematic reviews. *Sports Med*. 2018;48(1):189–205. <https://doi.org/10.1007/s40279-017-0781-4>. Medline:28887759
 28. Jamshidi L, Heyvaert M, Declercq L et coll. Methodological quality of meta-analyses of single-case experimental studies. *Res Dev Disabil*. 2018;79:97–115. <https://doi.org/10.1016/j.ridd.2017.12.016>. Medline:29289406
 29. Sorgente A, Pietrabissa G, Manzoni GM et coll. Web-based interventions for weight loss or weight loss maintenance in overweight and obese people: a systematic review of systematic reviews. *J Med Internet Res*. 2017;19(6):e229. <https://doi.org/10.2196/jmir.6972>. Medline:28652225
 30. Conley J, O'Brien CW, Leff BA et coll. Alternative strategies to inpatient hospitalization for acute medical conditions: a systematic review. *JAMA Intern Med*. 2016;176(11):1693–702. <https://doi.org/10.1001/jamainternmed.2016.5974>. Medline:27695822
 31. Corry M, While A, Neenan K et coll. A systematic review of systematic reviews on interventions for caregivers of people with chronic conditions. *J Adv Nurs*. 2015;71(4):718–34. <http://doi.org/10.1111/jan.12523>. Medline:25223528
 32. Rotta I, Salgado T, Silva M et coll. Effectiveness of clinical pharmacy services: an overview of systematic reviews (2000–2010). *Int J Clin Pharm*. 2015;37(5):687–97. <https://doi.org/10.1007/s11096-015-0137-9>. Medline:26001356
 33. Kitsiou S, Paré G, Jaana M. Systematic reviews and meta-analyses of home telemonitoring interventions for patients with chronic diseases: a critical assessment of their methodological quality. *J Med Internet Res*. 2013;15(7):e150–1. <http://doi.org/10.2196/jmir.2770>. Medline:23880072
 34. Anderson L, Taylor RS. Cardiac rehabilitation for people with heart disease: an overview of Cochrane systematic reviews. *Cochrane Database Syst Rev*. 2014;(12):CD011273. <https://doi.org/10.1002/14651858.CD011273.pub2>. Medline:25503364
 35. Sardanelli F, Bashir H, Berzaczky D et coll. The role of imaging specialists as authors of systematic reviews on diagnostic and interventional imaging and its impact on scientific quality: report from the EuroAIM Evidence-based Radiology Working Group. *Radiology*. 2014;272(2):533–40. <https://doi.org/10.1148/radiol.14131730>. Medline:24738613
 36. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. *Eur J Health Econ*. 2018;19(4):557–70. <https://doi.org/10.1007/s10198-017-0902-x>. Medline:28560520
 37. Salleh S, Thokala P, Brennan A et coll. Simulation modelling in healthcare: an umbrella review of systematic literature reviews. *Pharmacoeconomics*. 2017;35(9):937–49. <https://doi.org/10.1007/s40273-017-0523-3>. Medline:28560492
 38. Jimenez G, Virk AK, Tan WS et coll. Overview of systematic reviews of advance care planning: summary of evidence and global lessons. *J Pain Symptom Manage*. 2018;56(3):436–59.e25. <https://doi.org/10.1016/j.jpainsymman.2018.05.016>. Medline:29807158
 39. Labat T. Les effets de la cryothérapie sur les critères de jugement cliniques : une revue de revues systématiques. *Kinésithérapie Revue*. 2018;18(194):46. <https://doi.org/10.1016/j.kine.2017.12.015>.
 40. Alarcón MA, Diaz KT, Aranda L et coll. Use of biologic agents to promote bone formation in implant dentistry: a critical assessment

- of systematic reviews. *Int J Oral Maxillofac Implants*. 2017;32(2): 271–81. <http://doi.org/10.11607/jomi.5101>. Medline:27741328
41. Dosenovic S, Jelacic KA, Vucic K et coll. Comparison of methodological quality rating of systematic reviews on neuropathic pain using AMSTAR and R-AMSTAR. *BMC Med Res Methodol*. 2018;18(1):37. <https://doi.org/10.1186/s12874-018-0493-y>. Medline:29739339
 42. Gates A, Gates M, Duarte G et coll. Quality and risk of bias appraisals of systematic reviews are inconsistent across reviewers and centers. *J Clin Epidemiol*. 2020;125:9–15. <https://doi.org/10.1016/j.jclinepi.2020.04.026>. Medline:32416337
 43. Pieper D, Puljak L, Gonzalez-Lorenzo M et coll. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. *J Clin Epidemiol*. 2019;108:26–33. <https://doi.org/10.1016/j.jclinepi.2018.12.004>. Medline:30543911
 44. Lorenz RC, Matthias K, Pieper D et coll. A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. *Journal of Clin Epidemiol*. 2019;114:133–40. <https://doi.org/10.1016/j.jclinepi.2019.05.028>. Medline:31152864
 45. Kung J, Chiappelli F, Cajulis OO et coll. 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. Published 2010 Jul 16. <https://doi.org/10.2174/1874210601004020084>. Medline:31152864
 46. Whiting P, Savović J, Higgins JPT, et coll. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225–234
 47. 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>.
 48. Beaton DE, Bombardier C, Guillemin F et coll. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–91. <https://doi.org/10.1097/00007632-200012150-00014>. Medline:11124735
 49. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull*. 1971;76(5):378–82. <https://doi.org/10.1037/h0031619>
 50. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
 51. 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):633–39. <http://doi.org/10.1080/13561820600991181>. Medline:17095441
 52. Statistical Package for Social Sciences (SPSS), version 26.0. Chicago, IL : SPSS Inc; 2019.