

UNIVERSITÉ DE BOURGOGNE

UFR Médecine

THÈSE

Pour obtenir le grade de Docteur de l'Université de Bourgogne

Discipline : SCIENCES DE LA VIE

Par

Paul ORNETTI

Le 26 Octobre 2010

**VALIDATION TRANSCULTURELLE
DES CRITÈRES D'ÉVALUATION RAPPORTÉS PAR LE PATIENT
DANS L'ARTHROSE DES MEMBRES INFÉRIEURS**

Directeur de thèse :

Professeur Jean-Francis MAILLEFERT (Université de Bourgogne)

Jury :

Professeur Jean-Marie CASILLAS	Examineur	Université de Bourgogne
Professeur Francis GUILLEMIN	Examineur	Université de Nancy
Docteur Sébastien PARATTE	Examineur	Université d'Aix-Marseille
Professeur Christian TAVERNIER	Examineur	Université de Bourgogne
Docteur Véronique BONNIAUD	Rapporteur	Université de Franche-Comté
Professeur Serge POIRAUDEAU	Rapporteur	Université Paris Descartes

Résumé

Dans les pathologies ostéo-articulaires, les critères d'évaluation rapportés par les patients (*Patient-Reported Outcomes* ou PROs) sont désormais les critères de jugement incontournables, réclamés par les autorités sanitaires et les organismes de régulation. La validation de leurs propriétés psychométriques et leur adaptation transculturelle sont donc indispensables pour pouvoir les utiliser dans les essais thérapeutiques internationaux. Ce processus complexe de validation transculturelle permet d'élaborer une version du PRO dans le langage cible qui prenne en considération tous les spécificités socioculturelles et les équivalences linguistiques.

Ce travail de thèse a comme problématique la validation transculturelle des PROs spécifiquement développés pour les pathologies du genou et de la hanche. Après avoir explicité les recommandations publiées dans ce domaine puis avoir réalisé une revue de la littérature de la qualité méthodologique des auto-questionnaires disponibles pour ces pathologies articulaires, nous avons entrepris les validations transculturelles du questionnaire KOOS (ou *Knee injury and Osteoarthritis Outcome Score*), du questionnaire HOOS (ou *Hip disability and Osteoarthritis Outcome Score*) et de leurs formes réduites (KOOS-PS et HOOS-PS pour *Physical function Short form*). Ces auto-questionnaires, développés en langue anglaise au cours de la dernière décennie, permettent d'évaluer de manière globale (symptômes, retentissement fonctionnel dans les activités quotidiennes et sportives, qualité de vie) des pathologies du genou ou de la hanche, aussi bien traumatiques que dégénératives.

Après traduction et adaptation française selon les recommandations internationales, les versions françaises du KOOS, du HOOS, du KOOS-PS et du HOOS-PS ont ensuite été validées de manière prospective dans des cohortes françaises bi-centriques de patients arthrosiques suivis en milieu rhumatologique et orthopédique. Les résultats de ces travaux, publiés sous la forme de trois articles originaux, démontrent que les propriétés psychométriques des versions françaises de ces questionnaires sont satisfaisantes quelle que soit la sévérité de l'arthrose, en terme de faisabilité, de cohérence interne, de reproductibilité, de validité et de sensibilité au changement (après mise en place de prothèse totale ou viscosupplémentation par acide hyaluronique). Ces PROs capturent des concepts supplémentaires pertinents, comme le retentissement sur les activités sportives ou la qualité de vie, et permettent avec le même outil de mesure le suivi longitudinal de patients, depuis l'atteinte traumatique jusqu'à l'arthrose. Ils améliorent ainsi l'évaluation globale des pathologies du genou et de la hanche. La validation exhaustive, y compris transculturelle, permet à ce jour de considérer les questionnaires KOOS et le HOOS comme des PROs de référence. A ce titre, leur utilisation dans les essais thérapeutiques internationaux doit être encouragée.

Mots-clés : HOOS - KOOS - validation - adaptation transculturelle - arthrose - genou - hanche

Abstract

In joint diseases, *Patient-Reported Outcomes* (PROs) are now critical outcome measures, required by health authorities and regulatory agencies. Cross-cultural adaptation and validation of the psychometric properties of such PROs are needed in order to use them in international therapeutic trials. Thanks to this complex process of cross-cultural validation, the PRO in the target language can take into account the specific socio-cultural features and the linguistic equivalences. The aim of this thesis was to undertake the cross-cultural validation of PROs specifically developed for knee and hip disorders.

After explaining the published recommendations for cross-cultural adaptation and then conducting a thorough review of reports dealing with the methodological evaluation of self-administered questionnaires in these joint diseases, we carried out the cross-cultural validation of the KOOS questionnaire (*Knee injury and Osteoarthritis Outcome Score*), the HOOS questionnaire (*Hip disability and Osteoarthritis Outcome Score*) and their shortened versions (KOOS-PS and HOOS-PS for *Physical function Short form*). In the last ten years, these self-administered questionnaires were developed in English to assess the patients' opinion (symptoms, functional disability in activities of daily life or sports activities, quality of life) about their knee and hip problems, both for traumatic and degenerative joint diseases.

Following a translation and cross-cultural adaptation in French in accordance with international recommendations, the French versions of the KOOS, HOOS, KOOS-PS and HOOS-PS were validated in a bi-centric prospective study of osteoarthritis patients treated in two rheumatology and orthopedics departments. The results of this work, published as three original articles, demonstrate that the psychometric properties of the French versions of these questionnaires are satisfactory in terms of feasibility, internal consistency, reliability, validity and responsiveness (following total hip or knee replacement or intra-articular injection of hyaluronic acid), whatever the degree of severity of osteoarthritis. By capturing additional relevant concepts such as functional disability in sports activities or quality of life and allowing the longitudinal follow-up of patients (from traumatic lesions to osteoarthritis), with the same measurement tool, these PROs improve the overall assessment of hip and knee disorders. In the light of the exhaustive validation, including cross-cultural adaptation, the KOOS and HOOS questionnaires can now be considered as reference PROs in hip and knee disorders and their use in international therapeutic trials should be encouraged.

Keywords : HOOS - KOOS - validation - cross-cultural adaptation - osteoarthritis - knee - hip

TABLE DES MATIERES

INTRODUCTION	6
VALIDATION TRANSCULTURELLE DES AUTO-QUESTIONNAIRES : BASES FONDAMENTALES ET RECOMMANDATIONS	10
1. BASES FONDAMENTALES	11
2. TRADUCTION ET ADAPTATION TRANSCULTURELLE D'UN QUESTIONNAIRE	13
2.1 Etape préparatoire	13
2.3 Réunion de concertation.....	13
2.4 Rétro-traduction	13
2.5 Harmonisation et réunion d'experts.....	14
2.6 Pré-test	14
2.7 Validation finale.....	14
3. VALIDATION DE LA VERSION OBTENUE DANS LE NOUVEAU LANGAGE	15
3.1 Recommandations	15
3.2 Validité (validity).....	16
3.3 Cohérence interne (internal consistency).....	18
3.4 Fiabilité (reliability).....	18
3.5 Sensibilité au changement (responsiveness).....	19
3.6 Faisabilité (feasibility)	19
EVALUATION METHODOLOGIQUE DES AUTO-QUESTIONNAIRES DANS LES PATHOLOGIES DU GENOU ET DE LA HANCHE	21
1. PREAMBULE.....	22
2. METHODES	24
2.1 Recherche bibliographique	24
2.2 Critères d'inclusion et d'exclusion.....	24
2.3 Sélection des données : flow-chart	25
2.4 Présentation des résultats.....	26
2.5 Abréviations utilisées dans les fiches récapitulatives	27
3. AUTO-QUESTIONNAIRES SPECIFIQUES DU GENOU	29
3.1 Activity Rating Scale	29
3.2 American Academy of Orthopaedic Surgeons sport knee scales.....	30
3.3 Anterior Cruciate Ligament - Quality of Life Questionnaire	31
3.4 Cincinnati Knee Rating System subjective component	32
3.5 Edinburgh Knee Function Scale.....	33
3.6 Fulkerson-Shea patella-femoral joint evaluation score	34
3.7 Functional Index Questionnaire.....	35
3.8 Hughston subjective knee visual analog scale system (ou Flandry questionnaire)	36
3.9 International Knee Documentation Committee knee subjective form	37
3.10 Japanese Knee Osteoarthritis Measure.....	38
3.11 Knee outcome survey Activities of Daily Living Scale	39
3.12 Knee Quality of Life 26-item	40
3.13 Lequesne index knee osteoarthritis.....	41
3.14 Lysholm knee function scoring scale.....	42
3.15 Oxford knee questionnaire	43
3.16 Patello-femoral severity scale	44

4. AUTO-QUESTIONNAIRES SPECIFIQUES DE LA HANCHE.....	45
4.1 <i>Functional Recovery Score</i>	45
4.2 <i>Hip Outcome Score</i>	46
4.3 <i>Hip rating questionnaire</i>	47
4.4 <i>Lequesne index hip osteoarthritis</i>	48
4.5 <i>Lower Extremity Gain Scale</i>	49
4.6 <i>Lower Extremity Measure</i>	50
4.7 <i>Non-arthritic hip score</i>	51
4.8 <i>Oxford hip score</i>	52
4.9 <i>Patient Specific Index hip rating scale</i>	53
4.10 <i>Parkland and Palmer mobility score</i>	54
4.11 <i>Total Hip Arthroplasty Outcome questionnaire</i>	55
5. AUTO-QUESTIONNAIRES MIXTES (GENOU ET HANCHE) OU DES MEMBRES INFERIEURS.....	56
5.1 <i>American Academy of Orthopaedic Surgeons hip and knee scores</i>	56
5.2 <i>American Academy of Orthopaedic Surgeons lower limb score</i>	57
5.3 <i>High-Activity Arthroplasty Score</i>	58
5.4 <i>Intermittent And Constant OsteoArthritis Pain</i>	59
5.5 <i>Lower Extremity Activity Scale</i>	60
5.6 <i>Lower Extremity Functional Scale</i>	61
5.7 <i>New Zealand hip and knee score</i>	62
5.8 <i>OsteoArthritis Knee and Hip Quality Of Life</i>	63
5.9 <i>Reduced WOMAC function scale</i>	64
5.10 <i>Rheumatoid and Arthritis Outcome Score</i>	65
5.11 <i>WOMAC</i>	66
5.12 <i>WOMAC Function Subscale short form</i>	67
6. COMMENTAIRES	68
6.1 <i>Nombre de questionnaires croissant</i>	68
6.2 <i>Populations cibles et indications</i>	69
6.3 <i>Forme et contenu</i>	71
6.4 <i>Propriétés psychométriques</i>	72
6.5 <i>Retentissement sur les activités sportives</i>	74
6.6 <i>Retentissement sur la qualité de vie</i>	74
6.7 <i>Insuffisance de validation transculturelle</i>	75
VALIDATION TRANSCULTURELLE DES VERSIONS FRANÇAISES DES QUESTIONNAIRES	
KOOS, HOOS ET DE LEURS FORMES REDUITES.....	77
1. HISTORIQUE DU DEVELOPPEMENT DU KOOS ET DU HOOS.....	78
2. ADAPTATION TRANSCULTURELLE ET VALIDATION DE LA VERSION FRANÇAISE DU KOOS DANS LA	
GONARTHROSE.	81
2.1 <i>Résumé</i>	81
2.2 <i>Introduction</i>	82
2.3 <i>Méthodes</i>	84
2.4 <i>Résultats</i>	89
2.5 <i>Discussion</i>	97
3. ADAPTATION TRANSCULTURELLE ET VALIDATION DE LA VERSION FRANÇAISE DU HOOS DANS LA	
COXARTHROSE.....	101
3.1 <i>Résumé</i>	101
3.2 <i>Introduction</i>	102
3.3 <i>Méthodes</i>	104
3.4 <i>Résultats</i>	110

3.5 Discussion	120
4. PROPRIETES PSYCHOMETRIQUES DES VERSIONS FRANÇAISES DU KOOS ET DU HOOS REDUITS (KOOS-PS ET HOOS-PS)	124
4.1 Résumé	124
4.2 Introduction.....	125
4.3 Méthodes	127
4.4 Résultats	131
4.5 Discussion	135
5. SYNTHÈSE DE L'ÉVALUATION MÉTHODOLOGIQUE DES QUESTIONNAIRES KOOS ET HOOS	139
5.1 KOOS	139
5.2 KOOS-PS.....	140
5.3 HOOS.....	141
5.4 HOOS-PS	142
CONCLUSION	143
BIBLIOGRAPHIE.....	147
ANNEXES.....	161
ANNEXE 1. ARTICLE KOOS	162
ANNEXE 2. ARTICLE HOOS	169
ANNEXE 3. ARTICLE KOOS-PS/HOOS-PS	178
ANNEXE 4. VERSION FRANÇAISE DU KOOS	184
ANNEXE 5. VERSION FRANÇAISE DU HOOS	189

Introduction

Au cours des trente dernières années, de nombreux outils d'évaluation de l'état de santé des patients ont été développés dans les pathologies ostéo-articulaires afin d'appréhender leur retentissement global, tant en recherche clinique (1) qu'en pratique quotidienne (2). Certains de ces instruments sont destinés à apprécier des concepts aussi complexes et subjectifs que les symptômes, les capacités fonctionnelles ou la qualité de vie, en plus des données plus facilement mesurables comme celles de l'examen clinique, de la biologie ou de l'imagerie. Il est donc nécessaire d'utiliser des auto-questionnaires multidimensionnels pour évaluer au mieux les répercussions de ces maladies dans toute leur subjectivité et leur complexité. En rhumatologie, ces évaluations rapportées par le patient (PROs ou *Patient-Reported Outcomes* (3)), sont désormais reconnues comme des critères de jugement indispensables pour le suivi des patients dans les essais thérapeutiques (4-6). En effet, dans le domaine des pathologies rhumatismales, l'évaluation réalisée par le médecin sous-estime systématiquement le retentissement sur la fonction ou sur la qualité de vie (7, 8).

Cette multitude d'auto-questionnaires ne signifie pas pour autant ipso facto une amélioration de la qualité de l'évaluation des maladies ostéo-articulaires. En effet, beaucoup d'instruments de mesure ne bénéficient pas d'une validation suffisamment rigoureuse (9, 10). Par ailleurs, ce foisonnement de PROs conduit parfois à une perte d'information puisque les symptômes, les capacités fonctionnelles ou la qualité de vie sont évalués par des instruments différents en fonction du pays ou de la spécialité médico-chirurgicale. Ainsi, ce manque de consensus dans le choix des questionnaires ne permet pas la généralisation des résultats. Conscient de cette dérive progressive, certaines autorités sanitaires et agences de régulation comme la FDA (*Food and Drug Administration*) ou l'EMA (*European Medicine Agency*) exigent désormais que tous les instruments de mesure soient rigoureusement validés avant toute utilisation dans les essais cliniques (5, 6, 11).

Ces dernières années, dans le domaine des pathologies ostéo-articulaires, la validation des PROs est donc devenue l'un des axes de recherche majeurs pour les sociétés savantes comme l'OMERACT (*Outcomes Measures in Rheumatology Clinical Trials*) (12, 13). Des recommandations ((9, 10, 14-17) sont progressivement apparues afin de définir les critères permettant d'apprécier les qualités d'un questionnaire, souvent regroupées sous le terme de propriétés ou de qualités psychométriques. Ces recommandations ont permis d'importants progrès dans le domaine de l'évaluation des PROs. Ainsi, les auto-questionnaires les plus récents en Rhumatologie bénéficient indiscutablement d'une validation de bien meilleure qualité (10).

Outre ses propriétés psychométriques, un PRO doit avoir été adapté et validé dans de nombreux langages et cultures, afin de pouvoir être utilisé dans des essais internationaux de grande ampleur (18). Cette validation transculturelle doit être réalisée de manière extrêmement rigoureuse car les équivalences socioculturelles et les particularités linguistiques doivent être prises en considération, sous peine de perdre une partie des représentations et concepts compris dans l'auto-questionnaire (19). Des recommandations ont également vu le jour ces dernières années pour homogénéiser les pratiques dans le domaine de l'adaptation transculturelle des auto-questionnaires (20-22).

Concernant les pathologies du genou ou de la hanche, il est très difficile de choisir parmi les nombreux PROs développés spécifiquement pour ces deux articulations (23-25). Alors que la performance physique liée à ces atteintes articulaires est généralement mesurée par les mêmes tests (*performance-based outcome measures*), bien validés (test des 6 minutes, etc.) (26), les dimensions subjectives comme la sévérité symptomatique (douleur, raideur, etc.), l'impotence fonctionnelle, ou la qualité de vie sont évaluées par de nombreux auto-questionnaires disparates. Mis à part pour le WOMAC (*Western Ontario and Mc Master Universities index*) (27), l'utilisation d'un questionnaire commun à différentes spécialités

(orthopédie, rhumatologie, rééducation) demeure ainsi exceptionnelle. Par ailleurs, peu de PROs ont été validés aussi bien dans les pathologies traumatiques chez des patients actifs que dans les pathologies dégénératives chez des patients généralement plus âgés. De tels outils seraient pourtant particulièrement pertinents puisqu'ils permettraient le suivi longitudinal de pathologies aux liens évidents (de l'atteinte ménisco-ligamentaire traumatique à la gonarthrose) et faciliteraient la comparaison des résultats en recherche épidémiologique.

De ce postulat clinique est parti un vaste projet international visant à mettre au point un nouvel instrument de mesure pour le genou, le KOOS (ou *Knee injury and Osteoarthritis Outcome Score*) (28, 29) et un autre pour la hanche, le HOOS (ou *Hip disability and Osteoarthritis Outcome Score*) (30, 31). Ces auto-questionnaires ont la particularité d'intégrer le WOMAC dans sa forme complète mais comportent également des questions relatives à la fonction dans les activités sportives et à la qualité de vie. Le KOOS et le HOOS pourraient devenir des instruments de référence dans les essais cliniques internationaux sur les pathologies traumatiques et/ou dégénératives du genou ou de la hanche. A termes, ils pourraient s'intégrer dans d'éventuels scores composites de sévérité (32). Dans cette finalité, outre la validation de leurs propriétés psychométriques, ils doivent impérativement bénéficier d'un large programme d'adaptation transculturelle.

La problématique de ce travail de thèse s'articule ainsi autour de la validation française de ces questionnaires et de leurs formes réduites dans l'arthrose des membres inférieurs : KOOS pour le genou, HOOS pour la hanche, KOOS-PS et HOOS-PS. Dans une première partie seront définis les principes et recommandations dans le domaine de l'adaptation transculturelle. Nous ferons ensuite le point sur la qualité de l'évaluation méthodologique des auto-questionnaires existants pour le genou et/ou la hanche. Enfin, dans la troisième partie, nous exposerons les travaux ayant abouti à la validation des versions françaises des questionnaires KOOS, HOOS, KOOS-PS et HOOS-PS.

Validation transculturelle des auto-questionnaires : bases fondamentales et recommandations

1. Bases fondamentales

En médecine, la validation transculturelle d'un instrument de mesure est un processus indispensable pour permettre son utilisation à grande échelle dans les essais cliniques internationaux (17). L'intérêt grandissant pour la validation transculturelle des PROs se traduit par plus de 1000 articles publiés au cours de ces dix dernières années dans la base de données Medline, bien loin des premiers travaux confidentiels publiés par des psychologues cliniciens il ya trente ans (33). L'utilisation d'un même questionnaire validé dans différents langages présente des intérêts majeurs qui dépassent largement ceux liés au développement d'un nouvel instrument de mesure pour chaque langage cible (20, 34).

En pratique, la validation transculturelle d'un instrument de mesure est un processus bien plus complexe et long qu'il n'y paraît, qui requiert une méthodologie précise et rigoureuse. Les différentes étapes nécessaires à l'obtention d'une version validée dans un nouveau langage nécessitent la collaboration de nombreux intervenants : cliniciens de spécialités différentes, linguistes, épidémiologistes, statisticiens, sans oublier les patients. Avant de se lancer dans ce processus complexe, il faut veiller à ce que le questionnaire choisi n'ait jamais bénéficié d'une adaptation transculturelle dans le langage cible ou qu'un tel travail d'adaptation ne soit pas en cours. Il faut aussi s'assurer que la validation du questionnaire dans le langage source a été correctement réalisée. Dans l'immense majorité des cas, les auto-questionnaires sont initialement développés en langue anglo-saxonne (35). En fonction de la notoriété qu'ils acquièrent, des versions validées apparaissent progressivement dans des langues différentes, le plus souvent européennes et asiatiques. En revanche, peu de travaux de traduction et d'adaptation transculturelle ont été validés à l'heure actuelle en Afrique, Amérique Latine et en Océanie (36, 37).

L'adaptation transculturelle vise en premier lieu à s'assurer que la nouvelle version conserve tous les concepts inclus dans la première. En effet, dans les années 1990, des critiques ont émergé concernant la traduction littérale de certains questionnaires médicaux (38, 39) et la prise en compte insuffisante des équivalences culturelles et linguistiques (18, 40).

Afin d'homogénéiser les pratiques, des premières recommandations d'experts (14, 20, 38) ont progressivement vu le jour à cette époque et les définitions du type d'équivalences ont alors pu être précisées (19, 41). Des groupes de travail dédiés à la validation transculturelle ont été secondairement mis en place au sein de sociétés scientifiques réputées comme l'*International Society for Pharmacoeconomics and Outcome Research* (ISPOR) (22, 42), l'*International Quality of Life Assessment* (IQOLA) (38) ou l'*European Regulatory Issues and Quality of life Assessment* (ERIQA) (43), avec le soutien de la FDA (44) et de la *World Health Organization* (WHO) (45).

Dans une revue systématique de la littérature datant de 2008 (43), les différentes méthodes de traduction et d'adaptation transculturelle disponibles pour les questionnaires de qualité de vie liée à la santé ont été référencées : 33 articles ont été retrouvés au total par la recherche bibliographique, qui provenaient de 17 recommandations de groupes d'experts ou de sociétés savantes. A la lecture de ces différentes recommandations, on constate uniquement de petites différences entre elles, aussi bien dans la méthodologie employée que dans leur contenu, ainsi qu'un manque de consensus dans la terminologie de certains termes (*harmonisation*). Néanmoins, les différentes étapes qui permettent de mener à bien le processus de validation transculturelle restent communes à toutes ces recommandations, avec deux grandes parties bien distinctes : la traduction avec adaptation du questionnaire dans le langage cible puis la vérification des propriétés psychométriques de la version obtenue dans une population bien définie.

2. Traduction et adaptation transculturelle d'un questionnaire

(18-22, 34, 35, 46)

2.1 Etape préparatoire

Le coordonnateur du projet prend contact avec l'auteur du questionnaire source pour obtenir la permission de développer une nouvelle version dans le langage cible et lui proposer de participer à cette validation en tant qu'expert. Par ailleurs, le coordonnateur doit rechercher d'autres personnes susceptibles d'intégrer le groupe de travail durant les différentes étapes du processus, en expliquant son objectif et les différents concepts compris dans le questionnaire.

2.2 Traduction (forward translation)

Il s'agit de traduire le questionnaire du langage source dans le langage cible. Deux traducteurs au minimum sont nécessaires lors de cette étape. Ils doivent être bilingues mais de langue maternelle identique au langage cible. L'un des traducteurs doit connaître les concepts mesurés (médecin) mais l'autre pas (linguiste), pour éviter d'être tous les deux influencés par leur culture médicale et permettre si possible de détecter les termes ambigus pour les patients.

2.3 Réunion de concertation

Un comité comprenant les deux traducteurs, le coordonnateur et un méthodologiste expert dans le domaine de la validation transculturelle, natif du pays cible compare les traductions obtenues. Ils discutent ensemble des divergences obtenues et proposent des équivalences. L'auteur du questionnaire peut y participer pour clarifier certains points. A la fin de cette étape, une nouvelle version est adoptée par consensus dans le langage cible.

2.4 Rétro-translation

Des traducteurs bilingues (même nombre que pour la première traduction si possible), natifs du pays source et n'ayant ni participé à la première traduction, ni connaissance du questionnaire source participent à cette rétro-translation. Il s'agit de réaliser un contrôle qualité de la première traduction en amplifiant les imprécisions ou les inexactitudes dans la

formulation et en dépistant les questions problématiques. Afin d'éviter certains biais, les traducteurs ne doivent pas être des médecins ou des experts du domaine exploré.

2.5 Harmonisation et réunion d'experts

Toutes les traductions réalisées sont analysées lors d'une réunion comprenant toutes les personnes impliquées dans le processus d'adaptation transculturelle. L'objectif est que le questionnaire soit compréhensible par un enfant de 10-12 ans. Les équivalences sémantiques, idiomatiques, expérientielles, et conceptuelles sont discutées. D'éventuelles modifications sont entérinées pour aboutir à une nouvelle version du questionnaire.

2.6 Pré-test

Cette dernière version est testée parmi un échantillon représentatif de patients (au moins 10 personnes) ayant les mêmes caractéristiques que ceux de l'étude de validation ultérieure. Les patients font part aux experts de leurs suggestions sur la terminologie ou de leurs difficultés de compréhension, ce qui peut aboutir à certaines modifications.

2.7 Validation finale

Le coordonnateur du projet revoit la traduction obtenue, dans sa forme et dans son contenu, et édite la version finale dans le langage cible. Pour sa publication, il y ajoute une description des méthodes employées à chaque étape du processus de validation transculturelle.

Une fois la version traduite et adaptée dans le langage cible, il est obligatoire de vérifier ses qualités psychométriques dans la population cible. Cette validation doit être aussi rigoureuse que possible, selon les mêmes recommandations que pour tout nouveau questionnaire. En effet, les propriétés psychométriques du questionnaire traduit ne sont pas automatiquement superposables à celles du questionnaire source. Par ailleurs, grâce à cette nouvelle validation, la validité du questionnaire source est améliorée, tout comme sa notoriété.

3. Validation de la version obtenue dans le nouveau langage

3.1 Recommandations

Depuis une vingtaine d'années, des recommandations ont été proposées pour définir les critères d'évaluation des PROs dans le domaine de la santé (9, 10, 14-16, 47-49) dont le filtre OMERACT pour les pathologies ostéo-articulaire (49). Fin 2009, une mise à jour sur la validation des PROs a été publiée par la FDA dans le but d'homogénéiser les pratiques pour les essais thérapeutiques (5). En résumé, toutes ces recommandations proposent d'évaluer de manière systématique et détaillée les propriétés psychométriques suivantes : la validité, la fiabilité, la sensibilité au changement, en y ajoutant des données relatives à la faisabilité. L'importance de réaliser des adaptations transculturelles de ces PROs a également été soulignée pour faciliter la réalisation d'études internationales (48). Par ailleurs, le développement et la validation de formes réduites sont encouragées pour améliorer la faisabilité et la mise au point de critères composites (5).

Très récemment, à l'initiative d'un groupe de chercheurs hollandais (10), le projet international COSMIN (*COnsensus based Standards for the selection of health status Measurement INstruments* (51-53)) a vu le jour grâce à la collaboration d'épidémiologistes, de statisticiens, de psychologues et de cliniciens. Ce projet a pour but d'établir une check-list exhaustive de tous les critères nécessaires pour évaluer de manière rigoureuse les instruments de mesure de l'état de santé des populations, via l'avis d'experts, au moyen d'une méthode Delphi. Au terme de ce consensus d'experts, 10 critères comprenant 7 à 18 questions chacun ont finalement été retenus pour former la check-list (52) : la cohérence interne, la reproductibilité, l'erreur de mesure, les données manquantes, la validité de contenu, la validité de construit, les hypothèses formulées a priori, la validation transculturelle, la sensibilité de changement et l'interprétabilité. Cette check-list se présente sous la forme d'un questionnaire

de 80 questions au total, qui traduit l'exhaustivité de cette évaluation mais aussi sa lourdeur potentielle. Un score COSMIN permettant de coter l'évaluation méthodologique de chaque PRO est à termes envisagé (52). Avant de proposer une revue méthodologique dans la seconde partie de ce travail une revue des questionnaires disponibles dans les pathologies du genou et de la hanche, nous avons estimé nécessaire de définir tous les critères de la check-list COSMIN qui serviront de socle à cette revue.

3.2 Validité (*validity*)

La validité vise à vérifier si l'instrument mesure ce qu'il est censé mesurer (16). En pratique, plusieurs approches sont utilisées pour explorer les différentes facettes de la validité.

3.2.1 Validité d'apparence (*face validity*)

La validité d'apparence correspond à l'opinion de la valeur apparente que se fait l'utilisateur du questionnaire à la lecture de premier abord. Elle n'est pas mesurable statistiquement et est clairement insuffisante pour considérer qu'un questionnaire est valide, compte-tenu de son caractère superficiel et subjectif.

3.2.2 Validité de contenu (*content validity*)

Ce type de validité précise dans quelle mesure les questions sont pertinentes pour intégrer tous les concepts censés être mesurés par le questionnaire. Elle est généralement examinée de façon approfondie lors de la phase de développement du questionnaire (53). En pratique, aucun test statistique ne permet d'estimer la validité de contenu d'un questionnaire. L'interprétabilité des données exposées dans les résultats des études de validation (objectif du questionnaire, sélection des items pour chaque domaine, caractéristiques des populations, moyennes et dispersion des scores, comparaison des résultats entre différents échantillons) permet néanmoins d'estimer cette validité de contenu (16).

3.2.3 Validité de critère (*criterion validity*)

La validité de critère est plus simple à appréhender que la validité de contenu, pour peu qu'il existe un critère extérieur considéré comme un critère de référence (*gold standard*). L'amplitude des corrélations entre ce critère et les domaines du questionnaire définit ce type de validité. Toutefois, dans une large majorité de cas en rhumatologie, il n'existe pas de gold standard communément reconnu pour pouvoir estimer cette validité de critère. Il est alors nécessaire de bien préciser la validité de construit de l'instrument de mesure.

3.2.4 Validité de construit (*construct validity*)

La validité de construit explore les relations logiques devant exister entre le questionnaire et d'autres mesures habituelles de la pathologie évaluée. Cela justifie de poser des hypothèses a priori, de manière à préciser les validités convergente (corrélation supposée importante) et divergente (corrélation supposée faible) (54). Plus le nombre et la variété des corrélations recherchées sont importants, meilleure est la validité de construit du questionnaire. En revanche, si les hypothèses sont peu ou pas confirmées, la validité devient douteuse. La validité de construit est mesurée statistiquement par des tests de corrélation (Pearson ou Spearman).

3.2.5 Validité discriminante (*discriminant capacity*)

Elle correspond à la capacité du questionnaire à distinguer des échantillons différents au sein de la population étudiée (sains vs. malades ou atteinte modérée vs. atteinte sévère). La méthode la plus simple pour l'estimer est de comparer si les moyennes des scores des domaines sont statistiquement différentes entre les deux échantillons.

3.3 Cohérence interne (*internal consistency*)

Cette propriété, bien que décrite il y a plus de 50 ans (55) n'est devenue une propriété bien individualisée que depuis quelques années et son calcul est encore insuffisamment rapporté (10). La cohérence interne permet de vérifier la force des corrélations entre les items d'un domaine c'est-à-dire si les items de ce domaine mesurent tous le même concept (unidimensionnalité). Plus les items sont liés, plus la cohérence interne est importante. En termes de calcul statistique, le coefficient alpha de Cronbach permet de l'estimer.

3.4 Fiabilité (*reliability*)

Cette propriété psychométrique est l'un des critères essentiels de la validation d'un questionnaire car elle permet d'estimer les fluctuations de réponses, en l'absence de changement réel effectif. Pour les questionnaires, on utilise le plus souvent la reproductibilité dite en *test-retest* (9) (avec un intervalle de temps suffisant entre les deux mesures), en calculant le coefficient de corrélation intra-classe (ICC), avec une analyse de variance à un effet aléatoire. Plus l'ICC est proche de 1, plus la reproductibilité est bonne. Une variabilité trop importante contribue en revanche à diminuer la sensibilité au changement du questionnaire, en augmentant l'erreur de mesure. Il est donc essentiel d'évaluer cette erreur de mesure pour chaque questionnaire, en complément de l'ICC. Parmi les méthodes permettant son estimation, la méthode de Bland et Altman (56), sous la forme d'une représentation graphique (moyenne des évaluations en abscisse, différence entre les évaluations en ordonnée) est de plus en plus très utilisée. Elle permet de calculer la *Smallest Detectable Difference* (SDD), définie par $1.96 \times$ l'écart-type des différences entre les mesures répétées (57). Ce calcul du SDD doit être réalisé pour chaque domaine et dans chaque population étudiée car les variations observées peuvent être relativement importantes.

3.5 Sensibilité au changement (*responsiveness*)

La sensibilité au changement est une qualité psychométrique particulièrement importante dans les études thérapeutiques car plus l'instrument de mesure est sensible pour détecter un changement, plus la taille de l'échantillon peut être réduite. Parmi les méthodes de mesure de la sensibilité au changement (58), les plus utilisées pour les PROs sont le calcul de la taille de l'effet et de la réponse moyenne standardisée (RMS). En prenant N = nombre de sujets avant et après traitement et d la différence de score (avant – après traitement), on définit ces deux indices comme suit :

- Taille de l'effet = (moyenne des N différences d) / (écart-type des N scores avant traitement)
- RMS = (moyenne des N différences d) / (écart-type des N différences d).

3.6 Faisabilité (*feasibility*)

Même si elle n'est pas toujours considérée comme une propriété psychométrique à part entière (59), elle n'en demeure pas moins un critère d'évaluation à apprécier de manière systématique, ce qui est trop rarement le cas dans les études de validation de questionnaire (49). L'acceptabilité du questionnaire est perçue de façon subjective par le patient. D'autres mesures plus objectives permettent de juger de la faisabilité de l'outil de mesure.

3.6.1 Pourcentage de réponses manquantes (*missing data*)

Ce calcul simple est d'autant plus utile que le questionnaire comporte beaucoup de questions. Il permet parfois d'objectiver un problème lié à la présentation du questionnaire (lisibilité) ou à son contenu (absence de réponse car la patient n'est pas concerné par la question). Mais il n'est pas toujours rapporté dans les études de validation.

3.6.2 Effet plancher et effet plafond (*floor and ceiling effects*)

Les effets plancher ou plafond correspondent au pourcentage de patients qui atteignent le score minimum ou le score maximum pour un domaine du questionnaire étudié. Si ce pourcentage est trop important c'est qu'il y a un problème inhérent à la validité de contenu lors du processus de génération des items (10). Cela retentit sur la faisabilité du questionnaire mais aussi sur la capacité discriminante (les patients avec des scores extrêmes ne peuvent être distingués) et sur la sensibilité au changement (la taille de l'effet peut être réduite).

3.6.3 Temps de remplissage

Son impact sur la faisabilité est évident pour la pratique clinique. Il l'est aussi lors des essais thérapeutiques avec de volumineux cahiers d'observation contenant parfois plusieurs questionnaires. On considère généralement qu'en dessous de 10 questions, la faisabilité et l'acceptabilité sont bonnes. Au dessus de 30 questions, c'est-à-dire près de 15 minutes pour le remplir, la faisabilité devient médiocre. La lisibilité des questions, la clarté des consignes, le format du questionnaire peuvent également influencer sur ce temps de remplissage.

3.6.4 Cotation

La manière de coter varie selon les questionnaires, même si actuellement une échelle de Likert, avec un score défini pour chaque proposition de réponse est utilisée pour la plupart d'entre eux. La pondération de certaines réponses, la représentativité du score de chaque domaine dans le score total et l'éventuelle normalisation de ce dernier peuvent modifier sensiblement le temps de cotation. Dans le domaine des auto-questionnaires, très peu de scores correspondent à la simple somme des rangs des réponses, ce qui complique leur passage dans la pratique clinique (60).

Evaluation méthodologique des auto-questionnaires dans les pathologies du genou et de la hanche

1. Préambule

De nombreux PROs sont utilisés à l'heure actuelle pour évaluer les répercussions des pathologies articulaires, ce qui rend souvent difficile la comparaison des résultats entre les études utilisant des instruments de mesure différents. Même en se focalisant sur une articulation spécifique des membres inférieurs comme la hanche ou le genou et en se basant uniquement sur des auto-questionnaires, le clinicien ou le chercheur qui cherche à se faire une opinion parmi ce panel de questionnaires se trouve confronté à de sérieux écueils. La multiplication des revues de la littérature sur les échelles de mesures pour ces articulations en Chirurgie orthopédique (24, 61-65), en Rhumatologie (23, 25, 66, 67) ou en Médecine Physique (68-70) en est la meilleure illustration et traduit indubitablement une vraie demande de clarification. Bien que leurs approches ne soient pas toujours systématiques (24, 61, 63, 69), ces mises au point permettent de dresser un état des lieux des questionnaires disponibles tout en proposant un œil critique sur la qualité de leur validation. Cependant, les conclusions de ces articles sont en général rapidement obsolètes étant donné les publications permanentes dans ce domaine. Afin d'actualiser les données disponibles pour chaque PRO, des initiatives intéressantes ont été mises en place ces dernières années comme le projet PROQOLID (*Patient-Reported Outcome and Quality Of Life Instrument Database*), réalisé sous l'égide du *Mapi Research institute*. Ce projet international permet la diffusion sur internet (www.proqolid.org) des principales caractéristiques des PROs, sous forme d'une base de données régulièrement actualisée (71) qui reprend les principaux critères de validation de chaque instrument de mesure (plus de 600 instruments référencés et plus de 1000 traductions disponibles à l'heure actuelle).

Avant d'exposer, dans la dernière partie de ce travail de thèse, les données concernant la validation du KOOS et du HOOS, il nous est apparu indispensable de réaliser une revue méthodologique des questionnaires déjà disponibles pour le genou et la hanche. Outre la qualité de leur validation, cette recherche permettra de préciser leur spécificité concernant la maladie explorée et les domaines évalués mais aussi de comparer leur contenu (nombre de questions, cotation, etc.). La forme retenue pour la présentation des résultats n'a pas été celle d'une revue systématique et encore moins celle d'une méta-analyse pour pooler les données numériques par propriété psychométrique. En effet, du fait de leurs approches complexes, basées essentiellement sur le traitement statistique de données quantitatives, ces analyses systématiques, très en vogue en Rhumatologie, ne renseignent pas sur les données qualitatives des questionnaires (faisabilité, interprétabilité, contenu, traduction), ce qui occulte des aspects importants de la validation. Par souci didactique, nous avons préféré nous placer dans la position d'un médecin qui voudrait connaître les caractéristiques essentielles de chaque questionnaire disponible (contenu, validation transculturelle, propriétés psychométriques). Notre choix s'est alors naturellement porté sur une présentation sous forme de fiches récapitulatives, identiques pour chaque instrument, classées par localisation (genou, hanche). La détection des questionnaires a été effectuée grâce à une analyse bibliographique réalisée sur la base de données la plus consultée dans le monde médical : Medline.

Cette présentation, inhabituelle au premier abord pour une évaluation méthodologique, se rapproche du concept de check-list proposé dans le projet COSMIN (52) sans en avoir ni la lourdeur (80 questions pour la check-list contre seulement 18 critères par fiche) ni l'exhaustivité. Ce type de présentation permet de comparer visuellement des données des différents questionnaires avant d'aller rechercher si besoin des informations plus précises dans les articles de validation, référencés en haut de chaque fiche.

2. Méthodes

2.1 Recherche bibliographique

La base de données Medline a été consultée en utilisant l'équation de recherche suivante :

(hip OR knee) AND (outcome assessment OR questionnaire OR self assessment) AND (responsiveness OR validity OR reliability OR internal consistency OR translation OR cross cultural)

Il est admis que la première véritable évaluation d'un auto-questionnaire dans les pathologies articulaires date de 1977 (72). La recherche bibliographique a donc été restreinte à la période située entre janvier 1977 et janvier 2010. Les références citées dans les articles retenus qui n'avaient pas été détectées par l'équation de recherche ont été incluses dans l'analyse finale.

2.2 Critères d'inclusion et d'exclusion

Parmi les articles sélectionnés grâce à l'équation de recherche, seuls ont été conservés :

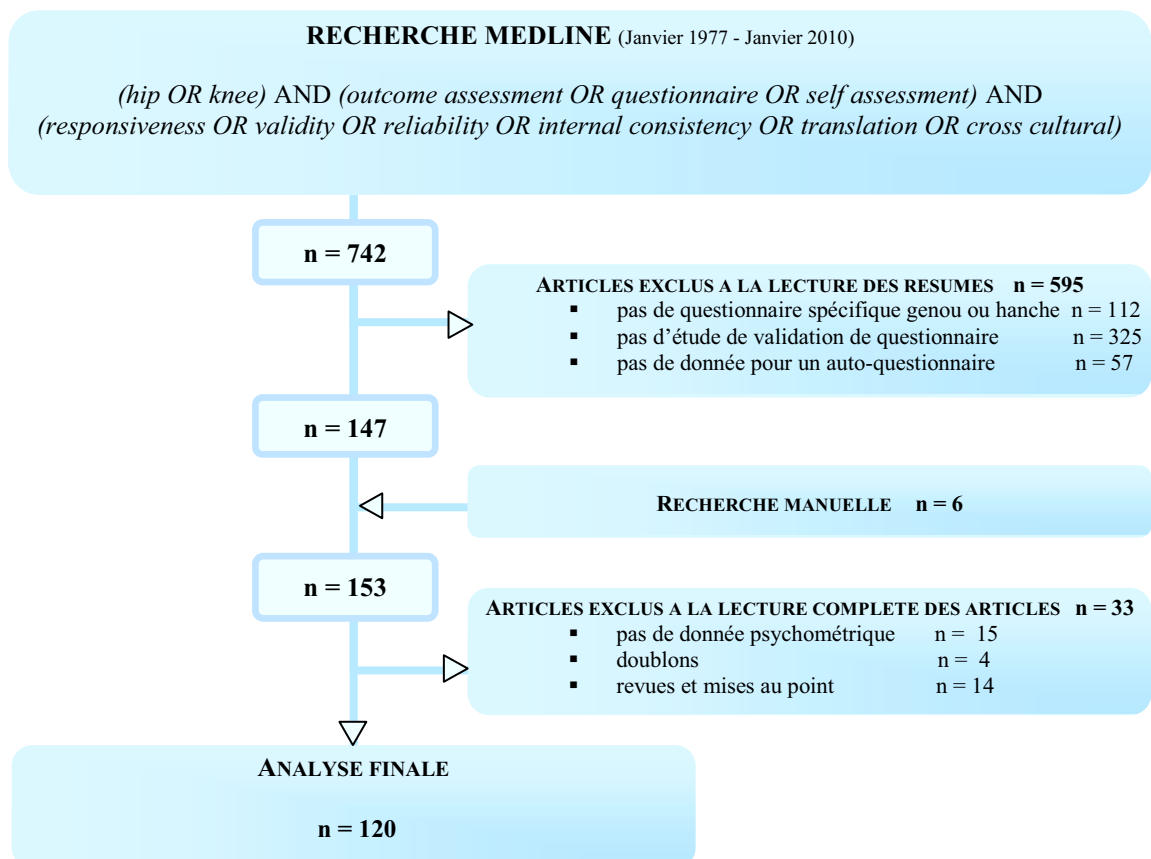
- les publications en anglais et en français, relatives à l'être humain
- dont l'objectif principal était la validation d'un auto-questionnaire
- spécifique de l'atteinte de la hanche et/ou du genou
- quelle que soit la pathologie sous-jacente concernée (arthrose, traumatisme, etc.).
- quel que soit le domaine évalué (symptômes, fonction, qualité de vie, activité, etc.)
- et qui rapportait au moins l'un des critères de jugement suivants selon la *Classical Test Theory* (73) : faisabilité, validité de construit, reproductibilité, cohérence interne, sensibilité au changement, effet plafond et plancher, erreur de mesure.

La validité de contenu n'a pas été incluse dans l'analyse finale car elle reste difficilement mesurable d'un point de vue statistique, ce qui est un obstacle à la présentation des résultats sous forme catégorielle (10).

A l'inverse, les articles présentant les caractéristiques suivantes ont été exclus :

- tous les articles de validation des questionnaires basés sur l'avis du médecin (*clinician-based outcome measure*) ou la performance physique (*performance-based outcomes measure*). Quand le questionnaire prenait en compte à la fois des critères-médecins et des critères-patients, comme cela peut être le cas pour certains questionnaires orthopédiques, les données issues des patients ont été colligées.
- tous les articles de validation de questionnaires non spécifiques de la hanche ou du genou, comme par exemple les questionnaires de qualité de vie génériques (questionnaires SF-36 (74), EuroQol-5D (75), etc.) ou spécifiques de maladies rhumatismales comme l'*Arthritis Impact Measurement Scales* (AIMS) (76).
- Tous les articles de validation de questionnaires constitués d'une seule question comme l'échelle d'activité de Tegner (77) ou la *Single Assessment Numeric Evaluation* (SANE) (78).

2.3 Sélection des données : flow-chart



2.4 Présentation des résultats

Les questionnaires retenus ont été classés en trois groupes (genou, hanche ou mixtes) et par ordre alphabétique. Une fiche récapitulative synthétique est associée à chaque questionnaire. La première moitié de la fiche comporte les données concernant le contenu du questionnaire : pathologie évaluée, validation initiale, validation transculturelle publiée, nombre et score des questions, nombre et dénomination des domaines, valeur et interprétation du score total. Dans la seconde moitié de la fiche apparaît l'évaluation des qualités psychométriques principales de chaque questionnaire.

- La faisabilité a été estimée par :

- le *pourcentage de réponses* au questionnaire : bon (+) si $< 2.5\%$ de réponses manquantes, moyen ($\pm$) si compris entre 2.5% et 5% et mauvais (-) si $> 5\%$ (16).
- la présence d'un *effet plancher ou plafond* : bon (+) si le score maximal ou minimal est atteint par $< 15\%$ des patients, mauvais (-) si atteint par $> 30\%$ des patients (10).
- le *temps de remplissage* : court (+) si le nombre de questions < 10 , modéré ($\pm$) entre 10 et 30 questions et long si > 30 questions.
- le *temps de cotation* : court (+) si le score total est la simple somme des scores par réponses, modéré ($\pm$) si la cotation nécessite une normalisation, long (-) si la cotation nécessite une pondération de certaines réponses ou des calculs supplémentaires.

- Afin de permettre une comparaison avec les données du KOOS (28) et du HOOS (30), les propriétés psychométriques ont été classées en insuffisante, modérée ou bonne (- / $\pm$ / +) selon les mêmes critères que ceux utilisés dans la validation française du KOOS (79) et du HOOS (80).

- la *reproductibilité* a été jugée bonne (+) si ICC > 0.8 , modérée ($\pm$) si ICC entre 0.7 et 0.8 et insuffisante (-) si < 0.7 (54).

- La *cohérence interne* a été jugée bonne (+) si le coefficient alpha de Cronbach > 0.8 , acceptable ($\pm$) si compris entre 0.7 et 0.8 et insuffisante (-) si < 0.7 (81).
 - La *validité de construit* a été jugée satisfaisante (+) si le coefficient de corrélation (Pearson ou Spearman) > 0.5 , modérée ($\pm$) entre 0.35 et 0.5 et insuffisante (-) si < 0.35 (82).
 - La *sensibilité au changement* a été considérée comme bonne (+) si la réponse moyenne standardisée ou la taille de l'effet étaient > 0.8 , modérée ($\pm$) entre 0.5 et 0.8 et faible (-) si < 0.5 (83).
- L'évaluation de l'erreur de mesure, quelle que soit la méthode pour l'estimer (SDD, coefficient de variation, etc.) a été aussi notée dans chaque fiche (évaluée oui/non).

2.5 Abréviations utilisées dans les fiches récapitulatives

- AAOS : American Academy of Orthopaedic Surgeons
- ADLS : Knee outcome survey Activities of Daily Living Scale
- AIMS : Arthritis Impact Measurement Scale
- EQ-5D : EuroQol-5 dimensions
- FIQ : Functional Index Questionnaire
- HADS : Hospital Anxiety and Depression Scale
- HAQ : Health Assessment Questionnaire
- HOS : Hip Outcome Score
- HSS : Hospital for Special Surgery knee scale

- ICOAP : Intermittent and Constant OsteoArthritis Pain
- IKS : International Knee Society score
- IKDC : International Knee Documentation Committee knee evaluation form
- KQoL-26 : Knee Quality of Life-26 item questionnaire
- LEAS : Lower Extremity Activity Scale
- LEGS : Lower Extremity Gain Scale
- LEM : Lower Extremity Measure
- LLFDI : Late-Life Function and Disability Instrument score
- MACTAR : McMaster Toronto ARthritis preference questionnaire
- MFA : Musculoskeletal Functional Assessment questionnaire
- OAKHQOL: OsteoArthritis Knee and Hip Quality Of Life questionnaire
- OARS : Older Americans Resources and Services function score
- POMS : Profile Of Mood States
- RAOS : Rheumatoid and Arthritis Outcome Score
- SF-36 : 36-Item Short Form survey instrument
- SF-12 : 12-Item Short Form survey instrument
- SIP : Sickness Impact profile

3. Auto-questionnaires spécifiques du genou

3.1 Activity Rating Scale

Activity Rating Scale (84)					
Etude de validation initiale		Marx RG et al (Am J Sports Med, 2001)			
Pathologie(s) cible(s)		Pathologies du genou (traumatiques ou dégénératives)			
Validation transculturelle publiée		Anglais			
Questions	Nombre	4			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction activités sportives			
Score total	Valeur	0 à 16			
	Interprétation	0 = retentissement maximal 16 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Echelle Tegner				
	Score Cincinnati				
	Echelle Daniel				
Sensibilité au changement					
Erreur de mesure					

3.2 American Academy of Orthopaedic Surgeons sport knee scales

American Academy of Orthopaedic Surgeons sports knee scales (85, 86)					
Etude de validation initiale		Johanson NA et al (J Bone Joint Surg Am, 2004)			
Pathologie(s) cible(s)		Pathologies traumatiques du genou			
Validation transculturelle publiée		Anglais			
Questions	Nombre	27			
	Score	0 à (5 ou 7)			
Domaines	Nombre	7			
	Dénomination	Score AAOS - Douleur - Limitation avant et après traumatisme - Activités sportives - Marche - Blocage			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	WOMAC				
	SF-36				
	Avis médecin				
Sensibilité au changement	Traitement orthopédique				
Erreur de mesure					

3.3 Anterior Cruciate Ligament - Quality of Life Questionnaire

Anterior Cruciate Ligament - Quality of Life Questionnaire (87, 88)					
Etude de validation initiale		Mohtadi N et al (Am J Sport Med, 1998)			
Pathologie(s) cible(s)		Pathologies ligament croisé antérieur			
Validation transculturelle publiée		Anglais			
Questions	Nombre	32			
	Score	0 à 100			
Domaines	Nombre	5			
	Dénomination	Symptômes - Activités professionnelles - Activités sportives - Style de vie - Retentissement social			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	EVA Douleur				
	Moments musculaires				
	Tests isométriques				
Sensibilité au changement	Chirurgie lig croisés				
Erreur de mesure					

3.4 Cincinnati Knee Rating System subjective component

Cincinnati Knee Rating System subjective component (89-92)					
Etude de validation initiale		Noyes FR et al (J Bone Joint Surg, 1983)			
Pathologie(s) cible(s)		Pathologies traumatiques du genou			
Validation transculturelle publiée		Anglais			
Questions	Nombre	6			
	Score	0 à 100			
Domaines	Nombre	1			
	Dénomination	Symptômes			
Score total	Valeur	0 à 20			
	Interprétation	20 = retentissement maximal 0 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Lysholm				
	Score AAOS				
	HSS				
	EVA douleur				
	SF-36				
Sensibilité au changement	Chir lig croisés				
	Chir cartilage				
Erreur de mesure					

3.5 Edinburgh Knee Function Scale

Edinburgh Knee Function Scale (93)					
Etude de validation initiale		Leigh Brown AP et al (Knee, 1999)			
Pathologie(s) cible(s)		Pathologies du genou sans indication chirurgicale			
Validation transculturelle publiée		Anglais			
Questions	Nombre	13			
	Score	0 à (3 à 5)			
Domaines	Nombre	3			
	Dénomination	Symptômes - Fonction - Activités de la vie quotidienne			
Score total	Valeur	0 - 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	HADS				
Sensibilité au changement					
Erreur de mesure					

3.6 Fulkerson-Shea patella-femoral joint evaluation score

Fulkerson-Shea patella-femoral joint evaluation score (94, 95)					
Etude de validation initiale		Fulkerson JP et al (J Bone Joint Surg, 1990)			
Pathologie(s) cible(s)		Pathologies fémoro-patellaires			
Validation transculturelle publiée		Anglais			
Questions	Nombre	7			
	Score	0 à (5 à 35)			
Domaines	Nombre	1			
	Dénomination	Symptômes			
Score total	Valeur	0 à 100			
	Interprétation	91-100 : excellent ; 81-90 : bon ; 71-80 : modéré ; <71 : insuffisant			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Echelle Tegner				
	Lysholm				
	Echelle Kujala				
	Score IKDC				
	MFA				
	SF-36				
Sensibilité au changement					
Erreur de mesure					

3.7 Functional Index Questionnaire

Functional Index Questionnaire (96, 97)					
Etude de validation initiale		Chesworth B et al (J of Orthop Res, 1989)			
Pathologie(s) cible(s)		Pathologies fémoro-patellaires			
Validation transculturelle publiée		Anglais			
Questions	Nombre	8			
	Score	0 à 2			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 16			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	EVA douleur				
	EVA fonction				
	Echelle Kujala				
	Test cinématique				
Sensibilité au changement	Rééducation				
Erreur de mesure					

3.8 Hughston subjective knee visual analog scale system (ou Flandry questionnaire)

Hughston subjective knee visual analog scale system (98, 99)					
Etude de validation initiale		Flandry F et al (Am J Sports Med, 1991)			
Pathologie(s) cible(s)		Pathologies du genou			
Validation transculturelle publiée		Anglais			
Questions	Nombre	28			
	Score	0 à 10			
Domaines	Nombre	1			
	Dénomination	Symptômes			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Lysholm				
	Score Cincinnati				
	Echelle Larson				
Sensibilité au changement					
Erreur de mesure					

3.9 International Knee Documentation Committee knee subjective form

International Knee Documentation Committee knee subjective form (89, 100-105)					
Etude de validation initiale		Irrgan J et al (Am J SportsMed, 2001)			
Pathologie(s) cible(s)		Pathologies du genou traumatiques et dégénératives			
Validation transculturelle publiée		Anglais - Allemand - Français - Hollandais - Italien - Espagnol - Brésilien - Thaïlandais			
Questions	Nombre	19			
	Score	1 à 11			
Domaines	Nombre	3			
	Dénomination	Symptômes - Activités sportives - Fonction			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Lysholm				
	EVA douleur				
	Ampl. articulaire				
	WOMAC				
	Oxford				
	KT-1000				
Sensibilité au changement	Chir lig croisés				
	Chir ménisques				
	Chir cartilage				
Erreur de mesure					

3.10 Japanese Knee Osteoarthritis Measure

Japanese Knee Osteoarthritis Measure (50)					
Etude de validation initiale		Akai M et al (J Rheumatol, 2005)			
Pathologie(s) cible(s)		Gonarthrose			
Validation transculturelle publiée		Japonais			
Questions	Nombre	25			
	Score	0 à (3 ou 4)			
Domaines	Nombre	3			
	Dénomination	Douleur - Activités quotidiennes - Retentissement social			
Score total	Valeur	0 à 92			
	Interprétation	0 = retentissement nul 92 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	WOMAC				
	EVA douleur				
Sensibilité au changement					
Erreur de mesure					

3.11 Knee outcome survey Activities of Daily Living Scale

Knee outcome survey Activities of Daily Living Scale (86, 106-109)					
Etude de validation initiale		Irrgang J et al (J Bone Joint Surgery, 1998)			
Pathologie(s) cible(s)		Pathologies du genou traumatiques et dégénératives			
Validation transculturelle publiée		Anglais - Allemand - Turc – Portugais			
Questions	Nombre	17			
	Score	0 à (4 ou 5)			
Domaines	Nombre	2			
	Dénomination	Symptômes – Fonction			
Score total	Valeur	0 à 80			
	Interprétation	0 = retentissement maximal 80 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Score Cincinnati				
	Lysholm				
	Score AAOS				
	Timed up test				
Sensibilité au changement	Rééducation				
Erreur de mesure					

3.12 Knee Quality of Life 26-item

Knee Quality of Life 26-item (110)					
Etude de validation initiale		Garratt A et al (Health Qual Life Outcomes, 2008)			
Pathologie(s) cible(s)		Pathologies ligamentaires et méniscales			
Validation transculturelle publiée		Anglais			
Questions	Nombre	26			
	Score	0 à 4			
Domaines	Nombre	3			
	Dénomination	Fonction - Limitations - Statut émotionnel			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	EQ-5D				
	SF-36				
	Lysholm				
	Arrêt de travail				
Sensibilité au changement	Chir ménisques				
Erreur de mesure					

3.13 Lequesne index knee osteoarthritis

Lequesne index knee osteoarthritis (111-121)					
Etude de validation initiale		Lequesne M et al (Scan J Rheum, 1987)			
Pathologie(s) cible(s)		Gonarthrose			
Validation transculturelle publiée		Français - Anglais - Allemand - Chinois - Coréen - Arabe			
Questions	Nombre	10			
	Score	0 à 8			
Domaines	Nombre	3			
	Dénomination	Douleur - Périmètre de marche - Activités quotidiennes			
Score total	Valeur	0 à 24			
	Interprétation	< 4 : retentissement minime ; 5 à 7 : modéré ; 8 à 10 : moyen ; 11 à 13 : important ; >13 : sévère			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	WOMAC				
	EVA douleur				
	Amplitude art				
	Avis patient				
	Avis médecin				
	Stade radio				
Sensibilité au changement	Prothèse totale				
	AINS				
	Rééducation				
Erreur de mesure					

3.14 Lysholm knee function scoring scale

Lysholm knee function scoring scale (77, 86, 89, 95, 110, 122-128)					
Etude de validation initiale		Lysholm J et al (Am J Sports Med, 1982)			
Pathologie(s) cible(s)		Pathologies du genou			
Validation transculturelle publiée		Anglais			
Questions	Nombre	8			
	Score	0 à 25			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 100			
	Interprétation	> 95 : excellent ; 84-95 : bon ; 65-83 : moyen ; < 65 : mauvais			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Score Cincinnati				
	Score AAOS				
	SF-36				
	WOMAC				
	Echelle Tegner				
	KOOS				
	KQoL-26				
	Score Marshall				
	Echelle Kujala				
	IKDC				
	Fulkerson-Shea				
Sensibilité au changement	Chir croisés				
	Rééducation				
	Chir ménisques				
	Chir cartilage				
Erreur de mesure					

3.15 Oxford knee questionnaire

Oxford knee questionnaire (101, 102, 129-135)					
Etude de validation initiale		Dawson et al (J bone Joint surg, 1998)			
Pathologie(s) cible(s)		Gonarthrose (avec indication de prothèse)			
Validation transculturelle publiée		Anglaise - Suédois - Allemand - Hollandais - Chinois - Italien - Thaïlandais			
Questions	Nombre	12			
	Score	1 à 5			
Domaines	Nombre	2			
	Dénomination	Douleur - Fonction			
Score total	Valeur	12 à 60			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	HAQ				
	SF-36				
	SIP				
	EVA douleur				
	Score AAOS				
	WOMAC				
	IKDC				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

3.16 Patello-femoral severity scale

Patello-femoral severity scale (136)					
Etude de validation initiale		Laprade JA et al (Clin Rehab, 2002)			
Pathologie(s) cible(s)		Pathologies fémoro-patellaires			
Validation transculturelle publiée		Anglais			
Questions	Nombre	10			
	Score	0 à 10			
Domaines	Nombre	1			
	Dénomination	Douleur			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	WOMAC				
	Echelle Hughston				
Sensibilité au changement					
Erreur de mesure					

4. Auto-questionnaires spécifiques de la hanche

4.1 Functional Recovery Score

Functional Recovery Score (137, 138)					
Etude de validation initiale		Zuckerman J et al (J Orthop Trauma, 2000)			
Pathologie(s) cible(s)		Après fracture hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	11			
	Score	0 à 4			
Domaines	Nombre	3			
	Dénomination	Activités quotidiennes basiques - Activités quotidiennes spécifiques - Mobilité			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Taux.mortalité				
	Réhospitalisation				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

4.2 Hip Outcome Score

Hip Outcome Score (139-142)					
Etude de validation initiale		Martin AL et al (Arthroscopy, 2006)			
Pathologie(s) cible(s)		Pathologies de hanche avec indication de prothèse			
Validation transculturelle publiée		Anglais			
Questions	Nombre	28			
	Score	0 à 4			
Domaines	Nombre	2			
	Dénomination	Activités de la vie quotidienne - Sports			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
Sensibilité au changement	Arthroscopie hanche				
Erreur de mesure					

4.3 Hip rating questionnaire

Hip rating questionnaire (143)					
Etude de validation initiale		Johanson NA et al (J bone Joint Surg, 1992)			
Pathologie(s) cible(s)		Pathologies de hanche avec indication chirurgicale			
Validation transculturelle publiée		Anglais			
Questions	Nombre	14			
	Score	1 à (3 à 25)			
Domaines	Nombre	4			
	Dénomination	Retentissement global - Douleur - Marche - Fonction			
Score total	Valeur	16 à 100			
	Interprétation	16 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Test 6 min				
	AIMS				
Sensibilité au changement	Prothèse totale				
	Ostéosynthèse				
	Hémi-prothèse				
Erreur de mesure					

4.4 Lequesne index hip osteoarthritis

Lequesne index hip osteoarthritis (111-121)					
Etude de validation initiale		Lequesne M et al (Scan J rheum, 1987)			
Pathologie(s) cible(s)		Coxarthrose			
Validation transculturelle publiée		Français - Anglais - Allemand - Chinois			
Questions	Nombre	10			
	Score	0 à 2 (sauf périmètre de 0 à 8)			
Domaines	Nombre	3			
	Dénomination	Douleur - Périmètre de marche - Activités quotidiennes			
Score total	Valeur	0 à 24			
	Interprétation	< 4 : retentissement minime ; 5 à 7 : modéré ; 8 à 10 : moyen ; 11 à 13 : important ; >13 : sévère			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	WOMAC				
	EVA douleur				
	Amplitude art				
	Avis patient				
	Avis médecin				
	Stade radio				
Sensibilité au changement	Prothèse totale				
	AINS				
	Rééducation				
Erreur de mesure					

4.5 Lower Extremity Gain Scale

Lower Extremity Gain Scale (144, 145)					
Etude de validation initiale		Zimmerman S et al (Arch Phys Med Rehabil, 2006)			
Pathologie(s) cible(s)		Fracture de hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	9			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Activités physiques			
Score total	Valeur	0 à 36			
	Interprétation	0 = retentissement maximal 36 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Retent. fonctionnel				
	Retent. social				
Sensibilité au changement					
Erreur de mesure					

4.6 Lower Extremity Measure

Lower Extremity Measure (146)					
Etude de validation initiale		Jaglal S et al (J bone Joint Surg AM (2000))			
Pathologie(s) cible(s)		Fracture de hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	29			
	Score	1 à 5			
Domaines	Nombre	1			
	Dénomination	Activités physiques			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Timed up and go				
	SF-36				
	Echelle OARS				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

4.7 Non-arthritic hip score

Non-arthritic hip score (147-149)					
Etude de validation initiale		Christensen C et al (Clin Orthop, 2003)			
Pathologie(s) cible(s)		Pathologies de hanche chez < 50 ans			
Validation transculturelle publiée		Anglais			
Questions	Nombre	20			
	Score	0 à 4			
Domaines	Nombre	4			
	Dénomination	Douleur - Symptômes - Fonction – Activités			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Harris				
	SF-12				
Sensibilité au changement	Arthroscopie hanche				
Erreur de Mesure					

4.8 Oxford hip score

Oxford hip score (150-156)					
Etude de validation initiale		Dawson J et al (J bone joint surg, 1996)			
Pathologie(s) cible(s)		Coxarthrose (avec indication de prothèse)			
Validation transculturelle publiée		Anglais - Hollandais - Allemand - Japonais - Français			
Questions	Nombre	12			
	Score	1 à 5			
Domaines	Nombre	2			
	Dénomination	Douleur - Fonction			
Score total	Valeur	12 à 60			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	AIMS				
	SF-36				
	Score Charnley				
	EQ-5D				
	Satisfaction patient				
	WOMAC				
	EVA douleur				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

4.9 Patient Specific Index hip rating scale

Patient Specific Index hip rating scale (58, 157, 158)					
Etude de validation initiale		Wright JG et al (J Bone Joint surg, 1994)			
Pathologie(s) cible(s)		Coxarthrose (avec indication de prothèse)			
Validation transculturelle publiée		Anglais			
Questions	Nombre	22			
	Score	1 à 7			
Domaines	Nombre	3			
	Dénomination	Symptômes - Activités vie quotidienne - Marche			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Harris				
	MACTAR				
	SF-36				
	WOMAC				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

4.10 Parkland and Palmer mobility score

Parkland and Palmer mobility score (159)					
Etude de validation initiale		Parker MJ et al (J Bone Joint Surgery, 1993)			
Pathologie(s) cible(s)		Fracture de hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	3			
	Score	0 à 3			
Domaines	Nombre	1			
	Dénomination	Mobilité			
Score total	Valeur	0 à 9			
	Interprétation	0 = retentissement maximal 9 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Mortalité à 1 an				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

4.11 Total Hip Arthroplasty Outcome questionnaire

Total Hip Arthroplasty Outcome questionnaire (160, 161)					
Etude de validation initiale		Liang MH et al (J Bone Joint Surg, 1991)			
Pathologie(s) cible(s)		Pathologies de hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	41			
	Score	Non spécifié			
Domaines	Nombre	7			
	Dénomination	Douleur - Activités professionnelles - Activités quotidiennes - Examen physique - Marche - Satisfaction - Complications			
Score total	Valeur	Non spécifiée			
	Interprétation	Non spécifiée			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	EVA douleur				
	Test de marche				
	SIP				
Sensibilité au changement					
Erreur de mesure					

5. Auto-questionnaires mixtes (genou et hanche) ou des membres inférieurs

5.1 American Academy of Orthopaedic Surgeons hip and knee scores

American Academy of Orthopaedic Surgeons hip and knee scores (85, 162)					
Etude de validation initiale		Johanson NA et al (J bone Joint Surg, 2004)			
Pathologie(s) cible(s)		Pathologies de genou et de hanche chez > 18 ans			
Validation transculturelle publiée		Anglais			
Questions	Nombre	16			
	Score	0 à (5 ou 7)			
Domaines	Nombre	3			
	Dénomination	Fonction - Douleur - Raideur et gonflement			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement minimal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plancher/plafond				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Douleur par médecin				
	Fonction par médecin				
	WOMAC				
	Oxford				
Sensibilité au changement	Chirurgies diverses				
Erreur de mesure					

5.2 American Academy of Orthopaedic Surgeons lower limb score

American Academy of Orthopaedic Surgeons lower limb scores (85)					
Etude de validation initiale		Johanson NA et al (J bone Joint Surg, 2004)			
Pathologie(s) cible(s)		Pathologies de genou et de hanche chez > 18 ans			
Validation transculturelle publiée		Anglais			
Questions	Nombre	7			
	Score	0 à (5 ou 7)			
Domaines	Nombre	3			
	Dénomination	Fonction - Douleur - Raideur et gonflement			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement minimal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plancher/plafond				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Douleur par médecin				
	Fonction par médecin				
	WOMAC				
	Oxford				
Sensibilité au changement					
Erreur de mesure					

5.3 High-Activity Arthroplasty Score

High-Activity Arthroplasty Score (163)					
Etude de validation initiale		Talbot S et al (J arthroplasty, 2010)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose avec indication de prothèse totale			
Validation transculturelle publiée		Anglais			
Questions	Nombre	4			
	Score	0 à (3 à 6)			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 18			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plancher/plafond				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Harris				
	Oxford				
	WOMAC réduit				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

5.4 Intermittent And Constant OsteoArthritis Pain

Intermittent And Constant OsteoArthritis Pain (164-167)					
Etude de validation initiale		Hawker GA et al (Osteoarthritis Cartilage, 2008)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose			
Validation transculturelle publiée		Anglais - Allemand - Hollandais - Tchèque - Italien- Espagnol - Norvégien - Suédois - Portugais			
Questions	Nombre	11			
	Score	0 à 4			
Domaines	Nombre	2			
	Dénomination	Douleur intermittente - Douleur permanente			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	KOOS				
	HOOS				
	WOMAC				
	Qualité de vie				
	LLFDI				
Sensibilité au changement	Prothèse totale				
Erreur de Mesure					

5.5 Lower Extremity Activity Scale

Lower Extremity Activity Scale (168, 169)					
Etude de validation initiale		Saleh K et al (J bone Joint Surg, 2005)			
Pathologie(s) cible(s)		Révision de prothèse de genou ou de hanche			
Validation transculturelle publiée		Anglais			
Questions	Nombre	18			
	Score	0 à 1			
Domaines	Nombre	1			
	Dénomination	Capacité de marche			
Score total	Valeur	0 à 18			
	Interprétation	0 = incapacité de marcher 18 = marche normale			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Podomètre				
	Niveau d'activités				
	WOMAC				
Sensibilité au changement	Prothèse totale				
Erreur de Mesure					

5.6 Lower Extremity Functional Scale

Lower Extremity Functional Scale (170-175)					
Etude de validation initiale		Binkley JM et al (Phys Ther, 1999)			
Pathologie(s) cible(s)		Pathologies des membres inférieurs			
Validation transculturelle publiée		Anglais			
Questions	Nombre	20			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 80			
	Interprétation	0 = retentissement nul 80 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	WOMAC				
Sensibilité au changement	Rééducation				
	Chir lig croisés				
	Prothèse totale				
Erreur de mesure					

5.7 New Zealand hip and knee score

New Zealand hip and knee score (176-178)					
Etude de validation initiale		Hadorn D et al (BMJ, 1997)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose avec indication de prothèse			
Validation transculturelle publiée		Anglais			
Questions	Nombre	8			
	Score	0 à (10 ou 20)			
Domaines	Nombre	4			
	Dénomination	Douleur - Fonction - Mobilité - Autres			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	WOMAC				
	EQ-5D				
	SF-12				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

5.8 OsteoArthritis Knee and Hip Quality Of Life

OsteoArthritis of Knee and Hip Quality Of Life (179, 180) = AMIQUAL en français (Arthrose des Membres Inférieurs et QUALité de vie)					
Etude de validation initiale		Rat AC et al (J clin Epidemiol, 2005)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose			
Validation transculturelle publiée		Français - Anglais			
Questions	Nombre	43			
	Score	0 à 10			
Domaines	Nombre	5			
	Dénomination	Activités physiques - Santé mentale - Douleur - Soutien social - Activités sociales			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement minimal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	WOMAC				
	Lequesne				
	Harris				
	KOOS				
	KOOS-PS/ HOOS-PS				
	Score IKS				
	EVA douleur				
	Périmètre de marche				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

5.9 Reduced WOMAC function scale

Reduced WOMAC function scale (181, 182)					
Etude de validation initiale		Whitehouse SL et al (J Bone Joint Surg, 2003)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose			
Validation transculturelle publiée		Anglais			
Questions	Nombre	7			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	WOMAC				
	Harris				
	SF-36				
	Oxford				
Sensibilité au changement	Prothèse totale				
Erreur de mesure					

5.10 Rheumatoid and Arthritis Outcome Score

Rheumatoid and Arthritis Outcome Score (183-185)					
Etude de validation initiale		Bremander AB et al (Health Qual Life Outcomes , 2003)			
Pathologie(s) cible(s)		Polyarthrite rhumatoïde			
Validation transculturelle publiée		Suédois - Anglais - Français - Turc			
Questions	Nombre	42			
	Score	0 à 4			
Domaines	Nombre	5			
	Dénomination	Douleur - Symptômes - Fonction activités quotidiennes - Fonction activités sports et loisirs - Qualité de vie			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	HAQ				
	AIMS-2				
Sensibilité au changement	Rééducation				
	Infiltr. corticoïdes				
Erreur de mesure					

5.11 WOMAC

WOMAC (27, 58, 113, 114, 117, 121, 175, 186-203)					
Etude de validation initiale		Bellamy N (J Rheumatol, 1988)			
Pathologie(s) cible(s)		Gonarthrose et coxarthrose			
Validation transculturelle publiée		Anglais - Suédois - Allemand - Coréen - Arabe – Chinois - Espagnol - Hébreu - Hollandais -Turc - Finlandais - Marocain			
Questions	Nombre	24			
	Score	0 à 4			
Domaines	Nombre	3			
	Dénomination	Douleur - Raideur - Fonction			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Ampl.articulaire				
	Lequesne				
	Lysholm				
	HAQ				
	EVA douleur				
	LEGS				
	Harris				
	AIMS				
	Oxford				
	Stade radiologique				
Sensibilité au changement	AINS				
	Prothèse				
	Rééducation				
	Viscosupplém.				
	Infiltr. corticoïdes				
	Anti-arthrosique				
Erreur de mesure					

5.12 WOMAC Function Subscale short form

WOMAC Function Subscale short form (204, 205)					
Etude de validation initiale		Tubach F et al (Ann Rheum Dis, 2005)			
Pathologie(s) cible(s)		Gonarthrose et Coxarthrose			
Validation transculturelle publiée		Anglais - Français			
Questions	Nombre	8			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 32			
	Interprétation	0 = retentissement nul 32 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	WOMAC				
	EVA douleur				
	EVA fonction				
	EVA app globale				
Sensibilité au changement	AINS				
Erreur de mesure					

6. Commentaires

La comparaison de la qualité de validation de PROs reste inévitablement sujette à discussion compte-tenu des populations cibles différentes, des domaines évalués distincts et de la méthodologie retenue (51). La revue présentée précédemment a néanmoins le mérite d'exposer de manière homogène les principales données disponibles concernant la validation de chaque auto-questionnaire de genou ou de hanche. Elle facilite ainsi leur comparaison et permet de s'orienter parmi ces nombreux PROs référencés dans Medline (39 au total, sans compter le KOOS et le HOOS), aux acronymes parfois très proches (LEGS vs. LEAS, HOS vs. HOOS) mais aux contenus différents. La lecture de cette revue impose plusieurs commentaires, touchant aussi bien à la finalité des questionnaires (indication et population cible), à leur forme (questions, domaines, cotation), à leur validation transculturelle, qu'à l'évaluation de leurs propriétés psychométriques.

6.1 Nombre de questionnaires croissant

Sur une trentaine d'années, le nombre d'auto-questionnaires développés et validés dans un champ d'investigation pourtant restreint (pathologies de genou et de hanche) a progressé de façon exponentielle. Dans le même temps, le nombre d'hétéro-évaluation (questionnaire rempli par le médecin) a progressivement diminué. A l'évidence cela traduit l'importance donnée au fil du temps à l'avis du patient pour le suivi et l'évaluation des maladies rhumatismales et traumatiques (64, 206). L'évaluation de ces auto-questionnaires est de plus en plus rigoureuse. Les plus récemment développés ont ainsi bénéficié, à de rares exceptions près (163, 168), de validations relativement complètes, grâce probablement aux recommandations des autorités de régulation et des sociétés savantes. On pourrait toutefois s'interroger sur l'intérêt de développer de nouveaux questionnaires compte-tenu du risque de

redondance et des difficultés à comparer les résultats des différentes études. Cette création de nouveaux outils de mesure s'explique en partie par le degré de spécialisation requis désormais dans chaque spécialité médico-chirurgicale : les orthopédistes ont développé un questionnaire spécifique de la révision d'arthroplastie (168) ou un autre instrument spécifique de la qualité de vie dans les entorses du ligament croisé antérieur (110) ; les rhumatologues ont mis au point un auto-questionnaire spécifique d'une atteinte mono-articulaire du membre inférieur dans les rhumatismes inflammatoires (183) ou un score spécifique de qualité de vie dans l'arthrose des membres inférieurs (179). Cela traduit un certain dynamisme de ces spécialités, dont les progrès et innovations thérapeutiques imposent une évaluation toujours plus précises des pathologies phares (pathologies traumatiques du genou, arthrose). En pratique, cette multiplication et cette spécialisation des questionnaires peuvent compliquer le dialogue entre les différents acteurs de santé et constitue un frein pour les essais cliniques multidisciplinaires de grande ampleur (15, 23, 65). Seuls de rares questionnaires comme le WOMAC (27) ont acquis une notoriété transversale, au moins en recherche clinique.

6.2 Populations cibles et indications

Les questionnaires spécifiques du genou ont bénéficié de plus nombreux développements et validations que ceux spécifiques de la hanche, avec des études de meilleure qualité méthodologique, en particulier pour la validité de construit et la validation transculturelle. Parmi eux, il faut noter la présence de questionnaires dédiés à une région anatomique précise (rotule ou ligaments croisés) dont la spécificité des symptômes (instabilité pour la rotule et laxité pour les ligaments) justifient des questions ou des domaines spécifiques. Pour les questionnaires spécifiques de hanche, les populations étudiées sont dans la grande majorité des cas constituées de patients très invalidés, au stade de chirurgie prothétique ou après fracture de hanche. Malheureusement, il existe de clairs déficits méthodologiques dans

l'évaluation proposée (validité de construit et cohérence interne) pour la plupart d'entre eux. Seuls le score Oxford et l'indice de Lequesne pour la hanche peuvent être considérés comme des instruments ayant fait preuve d'une validation complète et avec des résultats satisfaisants. Parmi les questionnaires mixtes, on retrouve des auto-questionnaires ayant fait l'objet d'études de validation rigoureuse dans différentes populations de malades (AAOS, LEFS, ICOAP, OAKHQOL, RAOS, WOMAC). On peut tout de même s'interroger sur la validité interne de certains domaines car les questions concernant les symptômes ou la fonction sont les mêmes pour le genou et la hanche, alors qu'il existe à l'évidence des spécificités propres à chaque articulation (boiterie, blocage, raideur, etc.).

Deux grandes indications prédominent pour les auto-questionnaires de genou et/ou de hanche : le suivi de pathologies traumatiques, surtout au genou (sujets en activité, sportifs) et le suivi de pathologies dégénératives, essentiellement à une phase pré-chirurgicale (sujets âgés ou sévèrement invalidés). Peu de questionnaires ont été correctement validés dans l'une et l'autre de ces deux indications (WOMAC, Oxford, LEFS, AAOS), alors que ces pathologies peuvent correspondre aux stades successifs de l'évolution de la maladie articulaire (de la pathologie ménisco-ligamentaire traumatique à la gonarthrose).

Par ailleurs, pour une même pathologie, rares sont les instruments à avoir été validés dans des populations différentes (arthrose de sévérité différente par exemple) afin d'évaluer leur capacité discriminante. Il faut souligner, qu'à de rares exceptions près (WOMAC, Activity Rating Scale, AAOS), la plupart des questionnaires n'ont pas été validés dans les pathologies dégénératives chez les gens jeunes et actifs (<50 ans) ou à un stade précoce, moins symptomatique. Le questionnaire Edinburgh knee function scale (93) est spécifiquement dédié aux pathologies dégénératives à un stade précoce, mais il ne présente pas de garanties suffisantes au niveau de sa qualité de sa validation, ce qui limite son intérêt.

6.3 Forme et contenu

Peu de questionnaires (10 sur 39) ont moins de 10 questions, ce qui limite leur utilisation en pratique clinique. Toutefois, ce grand nombre de questions s'explique par la nécessité d'évaluer les différents aspects du retentissement de ces maladies invalidantes et par le caractère multidimensionnel de certains domaines (fonction, qualité de vie). Il est intéressant de constater l'apparition récente de formes réduites de certains auto-questionnaires évaluant uniquement la fonction (181, 204), encouragées par la FDA sous réserve qu'elles bénéficient d'une validation propre (5).

Même si les questions qui les composent peuvent être très différentes, les domaines explorés sont généralement communs entre les questionnaires (symptômes, fonction). La marche, la course ou la boiterie sont souvent individualisées dans les questions, ce qui témoigne des répercussions locomotrices importantes des pathologies du genou et de la hanche.

A la lecture des fiches, on constate une certaine homogénéité dans la cotation des scores totaux, avec le plus souvent nécessité d'une normalisation pour obtenir un score total sur 100. Mais ces scores peuvent prêter à confusion car les scores chirurgicaux considèrent le 0 comme le retentissement maximal à l'inverse de la plupart des scores rhumatologiques ou de rééducation. Par ailleurs, la manière de calculer et d'interpréter les scores reste insuffisamment décrite dans les articles de validation, ce qui ajoute à la confusion. Il est rare que le score total corresponde à la simple somme des scores par question, ce qui est un obstacle à leur utilisation en pratique quotidienne. Alors que beaucoup de ces PROs ont été développés à visée chirurgicale, peu d'entre eux proposent un score seuil définissant, comme l'indice de Lequesne (111), l'indication potentielle de recours à la prothèse ou comme le Lysholm (128), la réussite du traitement chirurgical. Cela peut s'expliquer par les difficultés méthodologiques inhérentes à l'estimation de ces seuils.

6.4 Propriétés psychométriques

Concernant les critères choisis pour la *faisabilité*, les articles de validation manquent souvent de précision. Pour 2/3 des questionnaires, le pourcentage de données manquantes n'est pas disponible. La répartition des scores extrêmes dans la population cible avec les effets plancher ou plafond est rapportée moins d'une fois sur deux.

La *reproductibilité* en « test-retest » est nettement mieux évaluée dans les articles, de manière homogène du point de vue statistique avec le calcul de l'ICC. Pour la majorité des questionnaires, elle a été jugée satisfaisante ou bonne. Toutefois l'intervalle de temps entre le *test* et le *retest* variait énormément d'une étude à l'autre (de 24h à 1 mois). Ceci traduit la difficulté de choix entre d'une part le risque de mémorisation éventuel et d'autre part le risque de déstabilisation de la maladie durant cette période.

La *cohérence interne* est une propriété psychométrique particulièrement importante. Elle est rapportée de plus en plus fréquemment pour la validation des auto-questionnaires. Lorsqu'elle a été étudiée, elle a été jugée majoritairement satisfaisante pour les domaines des questionnaires, témoignant d'une bonne homogénéité des domaines. Toutefois, certains coefficients alpha de Cronbach rapportés sont très proches de 1 et posent la question de la redondance de certaines questions. Cela résulterait d'un possible défaut lors du développement du questionnaire (55).

La *validité de construit* est la propriété la mieux étudiée. Toutefois le choix des instruments servant à examiner le degré de corrélation varie considérablement d'une étude à l'autre et d'un questionnaire à l'autre, ce qui ne facilite pas la comparaison entre les questionnaires. Il est recommandé de réaliser pour chaque domaine une étude de la validité convergente avec un outil de mesure générique et un autre spécifique de la maladie explorée (207), ce qui est loin d'être le cas dans les questionnaires de notre revue. Les scores de questionnaires orthopédiques sont souvent corrélés à des paramètres cliniques (amplitudes articulaires,

fonction musculaire ou périmètre de marche). Cependant, le degré de corrélation avec ces paramètres cliniques est la plupart du temps inférieur à leur degré de corrélation avec d'autres PROs. Il faut noter que les hypothèses formulées a priori concernant la validité convergente et divergente sont insuffisamment détaillées et ce, malgré les recommandations (51). Très peu d'études de validation font état des corrélations entre les domaines explorés (symptômes, fonction, etc.) et la sévérité structurale (radiologique en particulier) qui est un autre critère de jugement important pour le suivi des maladies articulaires (1). Toutefois, dans les pathologies chroniques comme l'arthrose, le niveau de corrélations radio-cliniques restent souvent modéré (208).

La *sensibilité au changement* est souvent rapportée du fait de l'importance de cette qualité psychométrique dans le design des études thérapeutiques. Dans de nombreux questionnaires sur l'arthrose, seule la chirurgie prothétique a été évaluée. Les résultats sont effectivement bons en termes de taille de l'effet compte-tenu de l'efficacité reconnue et indiscutable de ce type de traitement radical. Néanmoins, leur sensibilité au changement pour des traitements moins efficaces (anti-inflammatoires, rééducation, etc.) mériterait d'être plus souvent évaluée. Par ailleurs, malgré le fait que ces outils de mesure soient centrés sur une ou deux articulations, la sensibilité au changement pour les gestes locaux (infiltration, viscosupplémentation) n'a été qu'exceptionnellement rapportée (WOMAC, validation française du RAOS (185)). Cette insuffisance a d'ailleurs été soulignée par les conclusions récentes d'un groupe de travail de l'OMERACT en raison du développement actuel des traitements intra-articulaires (209). Quant aux rares questionnaires n'ayant pas à l'heure actuelle de données publiées concernant la sensibilité au changement, leur utilisation doit être pour le moment déconseillée.

6.5 Retentissement sur les activités sportives

Jusqu'au milieu des années 1990, peu de questionnaires prenaient en compte les activités sportives dans l'évaluation des patients souffrant du genou ou de la hanche (98, 210). Les attentes nouvelles des malades en terme de loisirs, la conservation de l'état de santé même à des âges avancés ont favorisé l'apparition de ce domaine dans les questionnaires les plus récents (HOS, Activity Rating Scale, AAOS, RAOS, IKDC), sous l'impulsion des sociétés de Médecine du Sport (211). Ce domaine assez spécifique des pathologies musculo-squelettiques fait désormais partie intégrante de l'évaluation des répercussions fonctionnelles. Ce domaine présente l'avantage d'être en général très sensible au changement, même chez des patients à des stades précoces de la maladie. Il constitue par ailleurs un niveau d'exigence élevée pour l'évaluation des résultats de la chirurgie. L'évaluation de la sensibilité au changement paraît donc particulièrement déterminante pour ce domaine.

6.6 Retentissement sur la qualité de vie

La dégradation de la qualité de vie dans les pathologies rhumatologiques ou orthopédiques reste encore insuffisamment explorée (206, 212, 213). L'utilisation de questionnaires génériques reste de mise dans ces pathologies articulaires avec le SF-36 (214) comme chef de file. Des questionnaires spécifiquement dédiés à la qualité de vie dans les pathologies articulaires comme l'AIMS (76) ont aussi été validés dans de nombreuses maladies rhumatismales. Cette spécificité augmente encore désormais avec le développement récent de questionnaires de qualité de vie non seulement spécifiques d'une pathologie (pathologie ligamentaire, arthrose) mais aussi spécifiques d'une articulation du membre inférieur comme le KQol-26 (genou) ou l'OAKHQOL (genou ou hanche). D'autres comportent des questions relatives à la qualité de vie, comme le RAOS pour les atteintes du membre inférieur secondaires à un rhumatisme inflammatoire.

6.7 Insuffisance de validation transculturelle

La grande majorité des questionnaires répertoriés dans notre revue n'ont bénéficié d'aucune validation en dehors de celle originale en langue anglo-saxonne, ce qui limite leur utilisation dans des études internationales. Le développement d'un questionnaire spécifique d'une langue en lieu et place de la validation transculturelle d'un questionnaire existant ne paraît pas une solution alternative crédible (comme le Japanese Knee Osteoarthritis Measure (50)). Les questionnaires ayant des versions validées et publiées dans de nombreuses langues sont le WOMAC, l'ADLS, le Lequesne, l'IKDC, l'Oxford, l'ICOAP et le RAOS. Certains questionnaires considérés comme des références comme le Lysholm manquent curieusement de versions validées dans d'autres langues. La validation transculturelle améliore la validation globale de l'instrument de mesure, en fournissant de nouvelles données pour la validité de construit (nouvelles corrélations) et la sensibilité au changement (traitements différents).

6.8 Erreur de mesure

L'estimation de l'erreur de mesure est devenue ces dernières années un point incontournable dans le suivi des pathologies rhumatismales, aussi bien au niveau des PROs que des données biologiques ou radiologiques. Dans nos fiches, il nous est apparu intéressant de distinguer cette notion de la reproductibilité. Au regard des résultats, et même si on constate une amélioration dans les articles les plus récents, l'estimation de l'erreur de mesure n'apparaît que trop rarement dans les études de validation de questionnaires de hanche et de genou, en particulier en chirurgie orthopédique. Le fait que la chirurgie prothétique ait des effets largement supérieurs à la variabilité de mesure pourrait contribuer à expliquer cette insuffisance. Néanmoins, il demeure indispensable d'évaluer précisément l'erreur de mesure, surtout pour des traitements avec des effets beaucoup plus limités. Par ailleurs, une variation de mesure supérieure à l'erreur de mesure ne signifie pas forcément qu'elle soit cliniquement

pertinente, mais cette notion récente (MCID, *Minimal Clinically Important Difference*) dont la méthode de calcul fait débat (215) n'était qu'exceptionnellement rapportée dans les études de validation détectées par notre analyse bibliographique.

Au total, à la lumière des résultats exposés dans ce chapitre, peu d'auto-questionnaires ont bénéficié d'une validation convenable. Certains questionnaires du fait de l'exhaustivité de leur évaluation, de la richesse des corrélations, de la diversité des populations et de traitements évalués s'imposent comme des PROs de référence : l'Oxford dans le domaine chirurgical, le Lysholm dans les pathologies ménisco-ligamentaires (mise à part une validation transculturelle insuffisante) et le WOMAC pour l'arthrose. Tous ces questionnaires ont rapidement été adoptés par la communauté scientifique internationale et ont bénéficié de nombreuses validations transculturelles. L'IKDC et l'indice de Lequesne, très utilisés, pâtissent de leur manque de sensibilité au changement et de leur reproductibilité insuffisante.

Un auto-questionnaire de genou et/ou de hanche qui évaluerait le retentissement fonctionnel global depuis le problème traumatique jusqu'à un stade avancé d'arthrose serait particulièrement pertinent. Les conclusions de cette revue permettent de dresser la liste des critères qu'il devrait remplir pour devenir à son tour un PRO de référence :

1. *Validation de l'ensemble de ses propriétés psychométriques*
2. *Validation parmi des populations différentes, à des stades différents de la maladie*
3. *Validations transculturelles nombreuses et publiées*
4. *Validation d'une forme réduite pour des raisons de faisabilité*
5. *Validé de construit avec les questionnaires de référence de la maladie explorée*
6. *Evaluation distincte de la hanche et du genou, compte tenu des spécificités propres*
7. *Evaluation des symptômes et du retentissement fonctionnel de manière distincte*
8. *Evaluation pour des activités sollicitatrices comme les activités sportives*
9. *Evaluation de la qualité de vie*
10. *Evaluation de l'erreur de mesure pour chaque domaine du questionnaire*

**Validation transculturelle des versions
françaises des questionnaires KOOS,
HOOS et de leurs formes réduites**

Après avoir précisé les principes de la validation transculturelle puis fait le point sur la qualité de l'évaluation des PROs spécifiques du genou et/ou de la hanche, nous détaillerons dans cette dernière partie la validation française des questionnaires KOOS et HOOS. Ceci permettra de vérifier que ces deux questionnaires remplissent l'ensemble des critères exposés à la fin du chapitre précédent et sont à même de devenir des instruments de mesure de référence dans l'arthrose des membres inférieurs.

1. HISTORIQUE DU DEVELOPPEMENT DU KOOS ET DU HOOS

A la fin des années 1990, lors du traitement des données d'une cohorte suédoise de patients opérés pour une ménisectomie (19 ans de suivi en moyenne) (216), des médecins suédois des hôpitaux universitaires de Lund se sont trouvés confrontés à de sérieuses difficultés pour évaluer correctement ces patients sur le moyen et long terme. Faute de trouver un PRO répondant à leurs besoins, ils ont entrepris de développer un nouvel instrument de mesure adapté pour les pathologies traumatiques et dégénératives. Ce projet multidisciplinaire ambitieux, coordonné par le Professeur Ewa Roos va ainsi aboutir à la création du KOOS et du HOOS, dont les méthodes sont exhaustivement décrites sur le site internet dédié (217). Nous ne reprendrons ici que les phases principales de leur développement. Ceci permettra de comprendre le cheminement nécessaire à leur constitution et aussi l'intérêt de poursuivre ce projet d'épidémiologie clinique, en proposant leurs adaptations françaises. Après avoir réalisé une revue de la littérature sur le sujet, l'équipe de Roos a donc décidé de développer un auto-questionnaire spécifique pour le genou (KOOS) et un autre pour la hanche (HOOS). Afin d'améliorer leur validité de contenu pour l'arthrose, il fut décidé d'intégrer dans ces questionnaires l'ensemble des questions du WOMAC (27), après accord obtenu auprès de son auteur (N.Bellamy). Un groupe d'experts suédois et américains comprenant des chirurgiens orthopédiques, des médecins rééducateurs et des patients a alors été constitué pour élaborer d'abord le KOOS puis le HOOS (28). Lors du processus d'élaboration des questionnaires,

cinq domaines, cotés de 0 (retentissement maximal) à 100 (retentissement nul), ont été retenus pour caractériser au mieux les répercussions de ces pathologies articulaires : *Douleur* (comprenant le WOMAC douleur), autres *Symptômes* (comprenant le WOMAC raideur), *Fonction dans les activités quotidiennes* (identique au WOMAC fonction), *Fonction dans les activités sportives et de loisirs* et *Qualité de vie*. Les questions en rapport avec ces deux derniers domaines absents du WOMAC ont été générées à partir de questionnaires utilisant déjà ces concepts essentiels (87, 98). La finalité était d'aboutir à des questionnaires plus sensibles et discriminants que le WOMAC, en ajoutant des activités sollicitatrices sur la fonction physique et en faisant apparaître des notions nouvelles comme la qualité de vie.

Pour le KOOS, la revue de la littérature, l'avis des experts et une étude pilote sur des patients souffrant d'une arthrose post-traumatique ont permis de retenir au final 42 questions (28). Des versions anglaise et suédoise du KOOS ont été ensuite validées dans deux populations différentes, de manière quasi simultanée, en respectant pour la version suédoise les recommandations en vigueur pour l'adaptation transculturelle (18, 20). Les propriétés psychométriques de ces deux versions ont été considérées comme satisfaisantes en terme de reproductibilité, de cohérence interne, de validité convergente et divergente avec le SF-36 et de sensibilité au changement après chirurgie des ligaments croisés ou des ménisques sous arthroscopie (28, 29). Par la suite, ces résultats ont été confirmés dans d'autres études, qui ont apportés de nouvelles données rassurantes en terme de validité externe (Lysholm, EVA douleur, EQ-5D), de sensibilité au changement (prothèse totale de genou) et de capacité discriminante supérieure au WOMAC (arthrose modérée vs. arthrose évoluée) (218).

Pour le HOOS, les versions anglaise et suédoise initiales ont été développées par la même équipe pluridisciplinaire en utilisant le même processus de génération d'items puis de validation (30, 31). Compte-tenu des spécificités de l'articulation de la hanche, certaines questions (en particulier celles concernant le domaine *Symptômes*) sont très différentes de

celles du KOOS. Le HOOS comprend deux questions de moins que le KOOS, soit 40 questions au total. Ses bons résultats (30, 31) en termes de reproductibilité, de cohérence interne, de validité de construit (SF-36), de sensibilité au changement ont également été confirmés par des études de validation ultérieures (156, 219). En termes de faisabilité, ces deux questionnaires sont relativement longs à remplir, ce qui pouvait constituer un frein à leur utilisation, même dans les essais cliniques. Une forme réduite de chaque questionnaire a donc été développée en accord avec le Professeur Roos (220, 221) sous l'égide de l'OMERACT et de l'OARSI (222). Seules les questions des domaines du retentissement fonctionnel (pour les activités quotidiennes et sportives) ont été incluses dans un modèle de Rasch (223), grâce aux données de plus de 2000 patients. Cette méthode statistique a finalement permis d'aboutir au KOOS-PS et au HOOS-PS qui comprennent respectivement 7 et 5 questions (220, 221). La cohérence interne, la validité de construit et la sensibilité au changement après prothèse de genou ou de hanche ont été évaluées secondairement dans une étude canadienne (224). Le KOOS-PS et le HOOS-PS sont actuellement utilisés dans une large cohorte multicentrique concernant 11 pays. Cette étude vise à mettre au point un score composite comprenant en plus de la fonction, la douleur et la sévérité structurale, pour le recours à la prothèse totale dans la gonarthrose et la coxarthrose (32).

C'est donc le simple constat, par des cliniciens suédois, des difficultés à évaluer sur le moyen et le long terme les patients souffrant de pathologies de genou et de hanche qui a permis d'aboutir au développement puis à la diffusion internationale du KOOS et du HOOS. Afin de poursuivre leur validation, nous avons entrepris, en accord avec Ewa Roos, leur adaptation transculturelle en français. Ces travaux ont abouti à trois articles originaux dans *Osteoarthritis and Cartilage* (Annexes 1 à 3). A la demande de l'Ecole Doctorale, ces trois articles ont été traduits en français. Ils constitueront les prochains paragraphes.

2. ADAPTATION TRANSCULTURELLE ET VALIDATION DE LA VERSION FRANÇAISE DU KOOS DANS LA GONARTHROSE.

2.1 Résumé

Objectif. Adapter le *Knee injury and Osteoarthritis Outcome Score* (KOOS) en français et évaluer ses propriétés psychométriques.

Méthodes. La version française du KOOS a été développée selon les recommandations internationales en utilisant la technique de traduction/rétro-traduction afin obtenir une validité de contenu adéquate. Les données du KOOS ont été obtenues chez des patients avec une gonarthrose symptomatique. La version française a été évaluée dans deux populations de gonarthrose différentes, l'une sans indication de prothèse (groupe *médecine*) et la deuxième avec indication de prothèse totale de genou (groupe *chirurgie*). Les propriétés psychométriques suivantes ont été évaluées : la faisabilité par le pourcentage de réponses et la recherche d'effet plafond ou plancher ; la cohérence interne par le calcul de coefficient alpha de Cronbach ; la validité de construit par la recherche de corrélation avec les différents domaines de l'AMIQUAL (Arthrose des Membres Inférieurs et Qualité de vie), en utilisant le coefficient de corrélation de Spearman et la capacité discriminante par la différence des scores entre les groupes *médecine* et *chirurgie* ; la reproductibilité par le calcul du coefficient de corrélation intra-classe (ICC) et par la méthode graphique de Bland et Altman ; la sensibilité au changement en utilisant les données obtenues avant et 3 mois après la pose de prothèse de genou, par le calcul de la réponse moyenne standardisée (RMS) et de la taille de l'effet.

Résultats. 37 patients ont été inclus dans le groupe *médecine* (68 % femmes, âge moyen = 70 ± 10 ans) et 30 dans le groupe *chirurgie* (73 % femmes, âge moyen = 71 ± 10 ans). Le pourcentage de réponse a été excellent. Aucun effet plafond ou plancher n'a été observé, sauf pour le domaine *fonction sports et loisirs* (20.6 % des patients atteignaient le score le plus bas

dans le groupe médecine, 40% dans le groupe chirurgie avant la pose de prothèse totale mais 0 % après celle-ci). Les résultats concernant la cohérence interne (coefficient alpha de Cronbach allant de 0.76 à 0.93) et la validité de construit (convergente et divergente) étaient satisfaisants. Les patients en attente de chirurgie présentaient des scores significativement plus bas pour tous les domaines. La reproductibilité de mesure des différents domaines du KOOS était bonne ou excellente, avec des ICC compris entre 0.755 et 0.914. La sensibilité au changement était élevée, avec une RMS comprise entre 0.89 et 1.93 et une taille de l'effet comprise entre 1.31 et 2.8.

Conclusion. La version française du KOOS est un instrument de mesure valide, reproductible et sensible au changement qui permet de capturer certains aspects spécifiques de l'incapacité fonctionnelle chez les patients souffrant de gonarthrose, à l'origine d'une dégradation de leur qualité de vie.

2.2 Introduction

L'arthrose est une maladie articulaire dégénérative, caractérisée par une destruction progressive du cartilage. Elle peut toucher plusieurs articulations, en particulier les articulations portantes comme le genou. La douleur et l'incapacité qui découlent de la gonarthrose ont des répercussions importantes sur la qualité de vie des patients (225, 226).

Plusieurs instruments de mesure peuvent être utilisés pour évaluer les patients atteints de gonarthrose (24, 27, 225, 227). En particulier, le *Western Ontario and McMaster Universities Index* (WOMAC) est un auto-questionnaire spécifique de l'arthrose, validé et largement utilisé, qui permet d'évaluer la douleur, la raideur et les limitations fonctionnelles induites par l'arthrose (27). Le *Knee injury and Osteoarthritis Outcome Score* (KOOS) a été développé comme une extension du WOMAC pour les patients jeunes et/ou actifs souffrant

de gonarthrose ou de traumatismes du genou (28, 29). La différence principale entre le WOMAC et le KOOS est l'ajout dans ce dernier de deux domaines supplémentaires, le domaine *fonction sports et loisirs (fonction SL)* et le domaine *qualité de vie*. Le KOOS a montré une sensibilité au changement supérieure à celle du WOMAC chez les patients jeunes ou encore actifs, atteints de gonarthrose (28, 228, 229). Plusieurs études suggèrent que la population âgée actuelle, y compris les patients avec une gonarthrose sévère nécessitant une prothèse totale de genou, conserve une activité physique plus importante que la même population dans les décennies précédentes. En Allemagne, 42% des patients avec gonarthrose poursuivaient des activités sportives au moment du remplacement prothétique et 34% cinq ans après la pose de la prothèse (230). De plus, après l'opération, certains patients effectuaient des activités sportives qu'ils étaient incapables de faire avant la chirurgie. Par exemple, très peu de patients étaient capables de faire de la randonnée avant la pose de prothèse mais presque 30% d'entre eux cinq ans après la chirurgie (230). Dans une autre étude réalisée chez des patients en attente de prothèse de genou (âge moyen = 71 ans), respectivement 51% et 91% des patients interrogés considéraient que les domaines *fonction SL* et *qualité de vie* du KOOS étaient extrêmement importants (231). Le KOOS peut aussi prendre en compte des informations supplémentaires chez les patients âgés atteints de gonarthrose. Le KOOS a ainsi été évalué chez les patients avec une gonarthrose, y compris ceux avec une maladie avancée. Le KOOS a démontré sa validité, sa reproductibilité, sa sensibilité au changement chez ces patients avec gonarthrose, avant et après prothèse totale de genou (218, 231).

Avec le nombre croissant d'essais multicentriques internationaux et la nécessité d'obtenir des résultats significatifs pour les études épidémiologiques et/ou thérapeutiques, la validation et l'adaptation transculturelle des instruments de mesure de l'état de santé est devenu indispensable. L'adaptation transculturelle peut nécessiter non seulement la traduction mais aussi l'ajustement de mots, des tournures et des expressions. Ce processus peut entraîner

des modifications substantielles de certaines questions pour capturer le concept originel dans sa globalité. Des versions validées du KOOS sont actuellement publiées et utilisables en anglais, suédois (29), danois, allemand (232), chinois (82) et de nombreux autres langages (217). L'objectif de cette étude était de réaliser l'adaptation transculturelle du KOOS en français et d'évaluer les qualités psychométriques de la version française, en termes de faisabilité, de validité de construit, de reproductibilité et de sensibilité au changement.

2.3 Méthodes

Processus d'adaptation transculturelle

L'adaptation transculturelle a été réalisée selon les recommandations internationales publiées (20, 21). Trois personnes d'origine française (deux rhumatologues et un professeur d'anglais) ont traduit en français la version anglaise du KOOS, de manière indépendante. Une version finale a été obtenue après consensus entre les trois traducteurs. La rétro-translation a été réalisée par une personne bilingue (PB), d'origine anglaise, ne connaissant pas la version originale en anglais du questionnaire. Un comité multidisciplinaire a été constitué pour s'assurer que la traduction était compréhensible et pour vérifier les équivalences transculturelles entre la version source et la version finale. Ce comité comprenait trois rhumatologues (PO, LG, JFM), un chirurgien orthopédique (YJ), un rhumatologue épidémiologiste spécialisé dans les adaptations transculturelles (FG), un rhumatologue à la retraite souffrant de gonarthrose (JS) et un professeur d'anglais d'origine française (PG). La version finale a été pré-testée chez 15 patients souffrant de gonarthrose. Les patients devaient signaler s'ils comprenaient bien toutes les questions et s'ils rencontraient des problèmes avec certaines formulations.

Evaluation des propriétés psychométriques du KOOS

Design de l'étude. Etude prospective bi-centrique

Patients

Deux populations de patients ont été évaluées. Le groupe *médecine* comprenait des patients consultant pour une gonarthrose dans le service de Rhumatologie de l'hôpital universitaire de Dijon. Le groupe *chirurgie* comprenait des patients en attente de prothèse totale de genou dans le service de Chirurgie orthopédique des hôpitaux universitaires de Marseille. Les critères d'inclusion étaient un âge supérieur à 40 ans et une gonarthrose primitive selon les critères de l'*American College of Rheumatology* (233). Les patients du groupe *médecine* ne devaient pas avoir d'indication de prothèse totale de genou, contrairement aux patients du groupe *chirurgie*. Les patients devaient être capables de comprendre et de compléter seuls les auto-questionnaires. Les critères de non-inclusion étaient la présence d'une autre pathologie rhumatologique importante (lombalgies chroniques, arthrose d'une autre articulation des membres inférieurs, arthrose inflammatoire en poussée à l'examen clinique), l'utilisation de corticoïdes au cours des trois derniers mois et pour le groupe *médecine*, une modification du traitement de la gonarthrose dans les deux semaines à venir.

Questionnaires

Lors de l'évaluation initiale, les patients du groupe *médecine* devaient remplir la version française du KOOS et de l'AMIQUAL (Arthrose des Membres Inférieurs et QUALité de vie). L'AMIQUAL est un auto-questionnaire spécifique de qualité de vie dans la gonarthrose et dans la coxarthrose qui a été récemment validé (179). Il comprend 43 questions réparties en 5 domaines (*douleur, activités physiques, santé mentale, soutien social et activités sociales*) et 3 questions indépendantes (*activités sexuelles, activités professionnelles*). Les scores des domaines vont de 0 (retentissement maximal sur la qualité de vie) à 100 (retentissement nul).

Le KOOS comprend 5 domaines et 42 questions. La différence principale avec le WOMAC est l'ajout de deux domaines supplémentaires, un domaine comprenant 5 questions et évaluant le retentissement fonctionnel dans les activités sportives et de loisirs (*fonction SL*) et un domaine évaluant la qualité de vie sous forme de 4 questions. Par ailleurs, 4 questions ont été ajoutées au domaine *douleur* du WOMAC pour constituer le domaine *douleur* du KOOS et 5 questions ont été rajoutées au domaine *raideur* du WOMAC (qui comprend 2 questions) pour faire le domaine *symptômes* du KOOS. Les scores de chaque domaine sont ensuite normalisés sur une échelle allant de 0 à 100, la valeur 100 correspondant au meilleur résultat possible. Les patients du groupe *médecine* devaient remplir un second questionnaire KOOS deux semaines après et le renvoyer en utilisant une enveloppe préaffranchie. Les patients du groupe *chirurgie* devaient remplir le questionnaire KOOS avant et 3 mois après l'opération. Le traitement des données manquantes a été réalisé selon les recommandations issues du KOOS et de l'AMIQAL. Pour le KOOS, quand plus de deux réponses d'un domaine étaient manquantes, le score n'était pas calculé. Pour l'AMIQAL, quand plus de la moitié des réponses d'un domaine étaient manquantes, le score n'était pas calculé. Dans les autres situations, les données manquantes étaient remplacées par la moyenne des valeurs du domaine de chaque patient.

Analyse statistique

Faisabilité

La faisabilité a été évaluée par le pourcentage de réponses et la recherche d'un effet plafond ou plancher dans le groupe *médecine* et dans le groupe *chirurgie* (avant et trois mois après pose de prothèse totale de genou). Un effet plafond ou plancher été considéré comme présent si plus de 15% des patients atteignaient le score minimum ou maximum.

Cohérence interne

La cohérence interne a été évaluée par le calcul du coefficient alpha de Cronbach. Un coefficient de Cronbach supérieur à 0.7 est généralement considéré comme satisfaisant.

Validité de construit

La validité convergente et divergente a été déterminée en comparant les résultats du KOOS et ceux de l'AMIQUAL. Le coefficient de Spearman a été utilisé pour évaluer les corrélations entre leurs domaines respectifs. Des coefficients de corrélation >0.5 , entre $0.35-0.5$, et <0.35 ont été considérés respectivement comme des corrélations forte, modérée et faible (82). Des hypothèses ont été formulées a priori pour la validité convergente (corrélation modérée à forte) et divergente (faible corrélation), selon que les domaines mesurent des concepts proches ou éloignés et selon les données issues des précédentes études de validation du KOOS dans les autres langages (29, 228, 234). Il a été ainsi supposé que (1) les domaines *symptômes*, *douleur* et *fonction dans les activités quotidiennes (fonction AVQ)* seraient modérément ou fortement corrélés avec les domaines *douleur* et *activités physiques* de l'AMIQUAL et faiblement avec les autres domaines de l'AMIQUAL ; (2) Le domaine *fonction SL* du KOOS serait faiblement corrélé avec tous les domaines de l'AMIQUAL, comme ce domaine n'a démontré qu'une faible corrélation avec ceux du SF-36 (231) ; (3) Le domaine *qualité de vie* du KOOS serait corrélé modérément à fortement avec ceux de l'AMIQUAL.

Par ailleurs, les scores obtenus lors de la première évaluation dans les groupes *chirurgie* et *médecine* ont été comparés par ANOVA (après vérification de l'homogénéité des variances). Il a été supposé que les scores dans le groupe *chirurgie* devraient être significativement plus bas (retentissement plus important) que dans le groupe *médecine*.

Reproductibilité

La reproductibilité des domaines du KOOS a été évaluée à 15 jours d'intervalle dans le groupe *médecine*. Il a été admis que la probabilité d'un changement spontané d'état de la gonarthrose dans cet intervalle de temps était faible. L'évaluation de la reproductibilité a été réalisée par le calcul du coefficient de corrélation intra-classe (ICC), avec son intervalle de confiance à 95%. Un ICC supérieur à 0.80 est habituellement considéré comme le témoin d'une excellente reproductibilité. Par ailleurs, les représentations graphiques de Bland et Altman, dans laquelle la différence entre la première et la deuxième évaluation sont représentées en fonction de la moyenne des deux évaluations ont été obtenues (56). Ce type de représentation permet de décrire le pourcentage de patients et leur distribution à l'intérieur des bornes de l'intervalle de confiance à 95% pour la différence entre les mesures. La plus petite différence perceptible (SDD, *Smallest Detectable Difference*) qui correspond à 1.96 fois l'écart-type des différences entre les mesures répétées a ainsi pu être calculée. Le SDD correspond à la plus petite différence qui peut être distinguée de l'erreur de mesure.

Sensibilité au changement

La sensibilité au changement a été évaluée dans le groupe *chirurgie* en comparant les résultats avant et après prothèse. La réponse moyenne standardisée (RMS), c'est-à-dire la différence moyenne entre le score initial et le score 3 mois après la PTG, divisée par l'écart type de cette différence et la taille de l'effet à savoir, la différence moyenne entre le score initial et celui 3 mois après la PTG divisée par l'écart-type du score avant la chirurgie ont été calculées.

Le logiciel SPSS (*Statistical Package for the Social Sciences* (SPSS), version 14.0) a été utilisé pour l'analyse statistique. Le degré de significativité statistique était défini pour un $p < 0.05$.

2.4 Résultats

Seules de petites différences dans la structure des phrases des 42 questions ont été observées entre les versions anglaise et française, entre la version originale et rétro-traduite. Le comité multidisciplinaire discuta longuement la présentation du questionnaire, sa compréhensibilité et les différentes modalités de réponses avant d'obtenir un consensus sur la version finale.

Trente-sept patients ont été inclus dans le groupe *médecine* (âge moyen = 70 ± 10 ans, extrêmes 45-91 ans ; 68% femmes). Une large majorité d'entre eux (34/37) a retourné son second questionnaire 15 jours après, ce qui a permis d'évaluer la reproductibilité. Peu de réponses étaient manquantes (1.1%) et le score total a pu être calculé pour tous les domaines chez tous les patients. Un seul patient n'a pas rempli l'AMIQUAL. Parmi les 36 patients restants, plus de 99.5% des réponses étaient présentes. Les trois questions indépendantes de l'AMIQUAL, qui n'entrent pas en compte dans le calcul des scores des domaines, avaient un taux de non-réponses plus important : 88% de réponses manquantes pour la question 12 sur leurs activités professionnelles (la plupart des patients étant à la retraite) et 42 % de réponses manquantes pour les questions 22 et 23 sur l'activité sexuelle. Trente patients ont été inclus dans le groupe *chirurgie* (âge moyen 71 ± 10 ans, extrêmes 42-85 ans ; 73% femme). Chaque patient a rempli le questionnaire avant et trois mois après la chirurgie lors de la visite de contrôle. Parmi l'ensemble des questions, 4.7% des réponses étaient manquantes. Avant la chirurgie, les scores des domaines *symptômes* et *qualité de vie* ont été obtenus chez tous les patients. Le score *douleur* a été obtenu chez 29 patients, le score *fonction AVQ* a été obtenu chez 27 patients, le score *fonction SL* chez seulement 25 patients. Après chirurgie, le score *douleur* a été obtenu chez tous les patients alors que le score n'a pu être calculé chez respectivement un patient pour le domaine *symptômes* et le domaine *qualité de vie*, deux pour le domaine *fonction AVQ* et dix patients pour le domaine *fonction SL*. A la lecture des commentaires, il apparaissait que les réponses relatives au domaine *fonction SL* n'avaient pas

été remplies par certaines patients car ils n'avaient pas encore essayé de reprendre ce type d'activité (sauter, courir) au moment de l'évaluation post chirurgicale à 3 mois. Un score *douleur*, un score *symptômes* et trois scores *fonction AVQ* n'ont pas pu être calculés car les patients avaient oublié de remplir une page parmi les dix du questionnaire. Si ces pages oubliées n'avaient pas été prises en compte, le pourcentage de réponses manquantes serait passé à 3.7%. Aucun effet plancher ou plafond n'a été observé, sauf pour le domaine *fonction SL*. Pour ce domaine, un effet plafond a été observé dans le groupe *médecine* (20.6% des patients avec le score le plus bas). Dans le groupe *chirurgie*, l'effet plafond était plus prononcé avant la chirurgie (40% des patients avec le score le plus bas) mais disparaissait après la pose de PTG (0%). Les résultats pour la cohérence interne étaient bons avec un coefficient alpha de Cronbach compris entre 0.76 et 0.93 (tableau 1).

Domaines du KOOS (nombre de questions)	Coefficient alpha de Cronbach
<i>Douleur (9)</i>	0.84
<i>Symptômes (7)</i>	0.76
<i>Fonction AVQ (17)</i>	0.93
<i>Fonction SL (5)</i>	0.84
<i>Qualité de vie (4)</i>	0.83

Tableau 1. Cohérence interne des domaines du KOOS

Les résultats pour la validité convergente et divergente sont présentés dans le tableau 2.

KOOS AMIQUAL	<i>Douleur</i>	<i>Symptômes</i>	<i>Fonction</i> <i>AVQ</i>	<i>Fonction SL</i>	<i>QdV</i>
<i>Activités physiques</i>	0.45 (<i>p</i> = 0.008)	0.29 (<i>p</i> = 0.098)	0.65 (<i>p</i> < 0.001)	0.30 (<i>p</i> = 0.084)	0.53 (<i>p</i> = 0.001)
<i>Douleur</i>	0.42 (<i>p</i> = 0.014)	0.35 (<i>p</i> = 0.042)	0.48 (0.004)	0.20 (<i>p</i> = 0.258)	0.54 (<i>p</i> = 0.001)
<i>Santé mentale</i>	0.46 (<i>p</i> = 0.007)	0.47 (<i>p</i> = 0.005)	0.50 (<i>p</i> = 0.002)	0.14 (<i>p</i> = 0.44)	0.72 (<i>p</i> < 0.001)
<i>Soutien social</i>	0.04 (<i>p</i> = 0.82)	- 0.12 (<i>p</i> = 0.5)	0.07 (<i>p</i> = 0.68)	0.05 (<i>p</i> = 0.77)	0.08 (<i>p</i> = 0.66)
<i>Activités sociales</i>	0.30 (<i>p</i> = 0.089)	0.29 (<i>p</i> = 0.091)	0.35 (<i>p</i> = 0.043)	0.18 (<i>p</i> = 0.31)	0.58 (<i>p</i> < 0.001)

Tableau 2. Validité de construit : corrélations entre les domaines du KOOS et de l'AMIQUAL (coefficient de Spearman)

Comme supposé, des corrélations fortes ou modérées ont été observées entre les domaines du KOOS et de l'AMIQUAL mesurant les mêmes concepts : *fonction AVQ* du KOOS avec *activités physiques* de l'AMIQUAL ($r_s = 0.65$ ($p < 0.001$)), *douleur* du KOOS et *douleur* de l'AMIQUAL ($r_s = 0.42$ ($p = 0.01$)). Une corrélation modérée était aussi retrouvée comme attendu entre le domaine *fonction AVQ* du KOOS et le domaine *douleur* de l'AMIQUAL et entre le domaine *douleur* du KOOS et le domaine *activités physiques* de l'AMIQUAL. Comme l'AMIQUAL est un questionnaire de qualité de vie, il n'était pas surprenant d'observer une bonne corrélation entre le domaine *qualité de vie* du KOOS et tous les domaines de l'AMIQUAL, sauf pour le domaine *soutien social*. Les faibles corrélations entre

le domaine *fonction SL* du KOOS et tous les domaines de l'AMIQUAL avaient déjà été observées avec les domaines du SF-36 (231). Quelques résultats inattendus ont été obtenus, en particulier une corrélation modérée à forte entre le domaine *santé mentale* de l'AMIQUAL et tous les domaines du KOOS, sauf pour le domaine *fonction SL*.

Les résultats obtenus entre les groupes *médecine* et *chirurgie* sont présentés dans le tableau 3.

Domaines du KOOS	<i>Groupe médecine</i> moyenne (DS)	<i>Groupe chirurgie</i> moyenne (DS)	p
<i>Douleur</i>	51 (16)	44 (14)	0.04
<i>Symptômes</i>	59 (20)	49 (18)	0.02
<i>Fonction AVQ</i>	53 (16)	45 (14)	0.04
<i>Fonction SL</i>	24 (20)	11 (16)	0.01
<i>Qualité de vie</i>	37 (17)	23 (14)	0.001

Tableau 3. Comparaison entre les domaines du KOOS du groupe *médecine* sans indication de PTG (n = 37) and ceux du groupe *chirurgie* (n = 30) avec indication de PTG (ANOVA).

DS : déviation standard

Comme attendu, les patients en attente de chirurgie avaient des scores du KOOS significativement plus bas (retentissement plus important) pour chaque domaine.

La reproductibilité de mesure de tous les domaines du KOOS était bonne à excellente, avec des ICC compris entre 0.755 et 0.914 (tableau 4). Le SDD était compris entre 13.4 (domaine *douleur*) et 21.1 (domaine *qualité de vie*).

Domaines du KOOS	1^{ère} évaluation moyenne (DS)	2^{ème} évaluation moyenne (DS)	ICC (IC 95%)	SDD
<i>Douleur</i>	50 (15)	51 (13)	0.883 (0.780-0.940)	13.4
<i>Symptômes</i>	58 (19)	58 (17)	0.914 (0.834-0.956)	15.5
<i>Fonction AVQ</i>	51 (14)	50 (15)	0.859 (0.736-0.927)	15.4
<i>Fonction SL</i>	21 (17)	21 (15)	0.824 (0.675-0.908)	19.6
<i>QdV</i>	35 (15)	34 (15)	0.755 (0.563-0.87)	21.1

Tableau 4. Scores moyens du KOOS et reproductibilité de mesure à 15 jours d'intervalle

(n = 34)

SDD : plus petite difference perceptible (*smallest detectable difference*)

ICC : coefficient de corrélation intra-classe

IC 95% : intervalle de confiance

DS : déviation standard

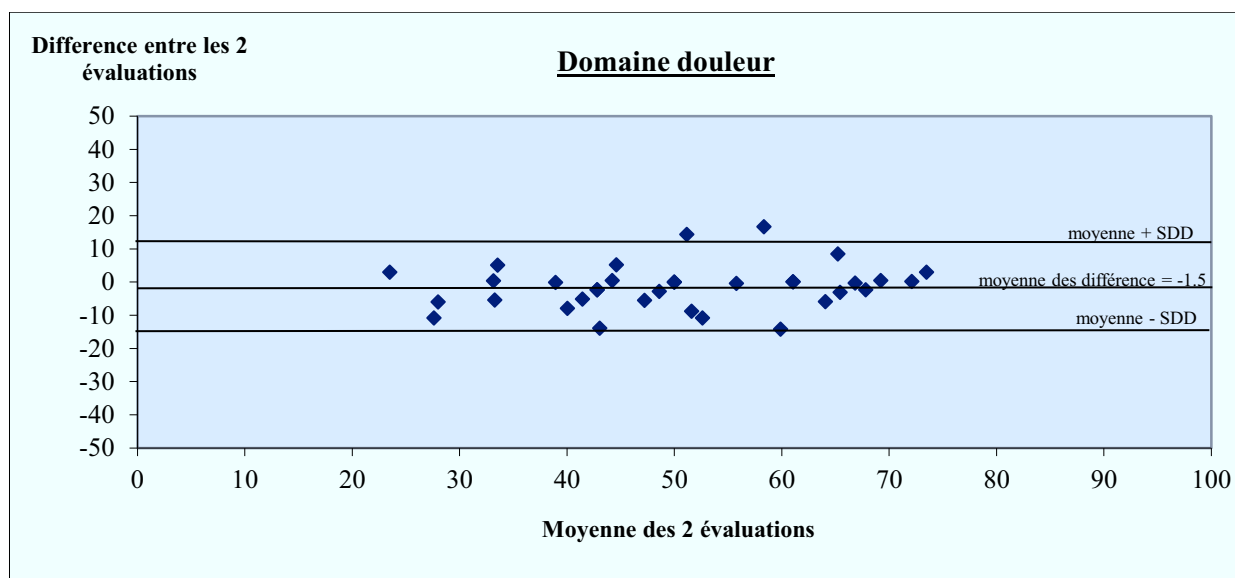
Les représentations graphiques de Bland et Altman sont présentées dans la figure 1 (page suivante). La différence entre les mesures répétées ne dépend pas de la moyenne des deux mesures pour aucun des domaines.

Figure 1.

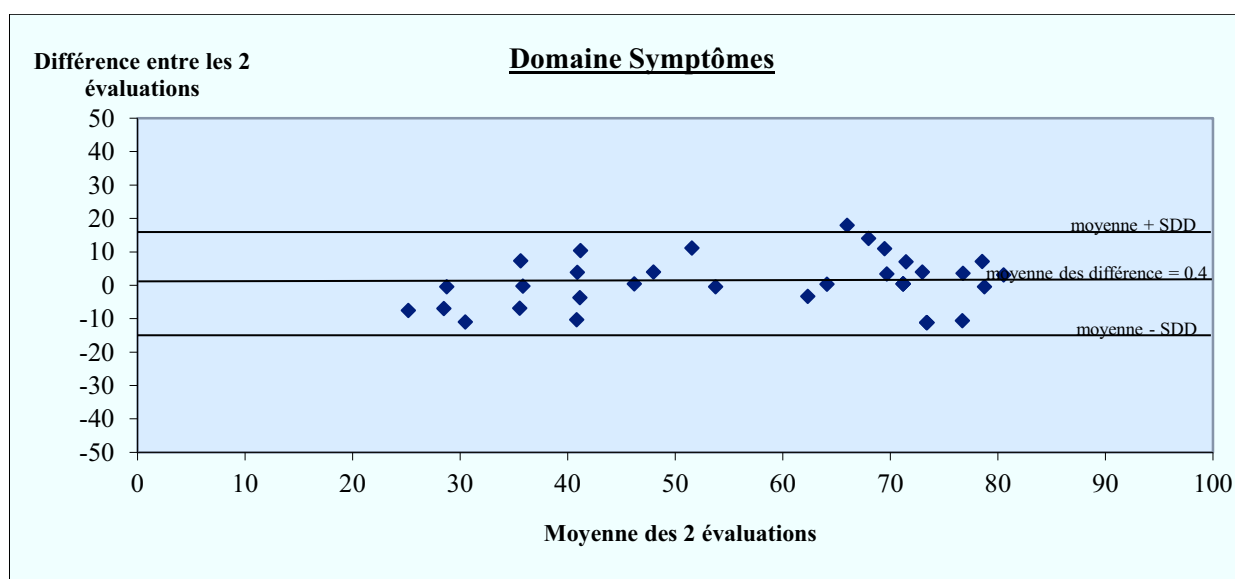
Reproductibilité de la version française du KOOS (méthode graphique de Bland et Altman). Deux évaluations successives ont été réalisées à 15 jours d'intervalle.

(a) domaine *douleur* **(b)** domaine *symptômes* **(c)** domaine *fonction AVQ* **(d)** domaine *fonction SL* **(e)** domaine *qualité de vie*

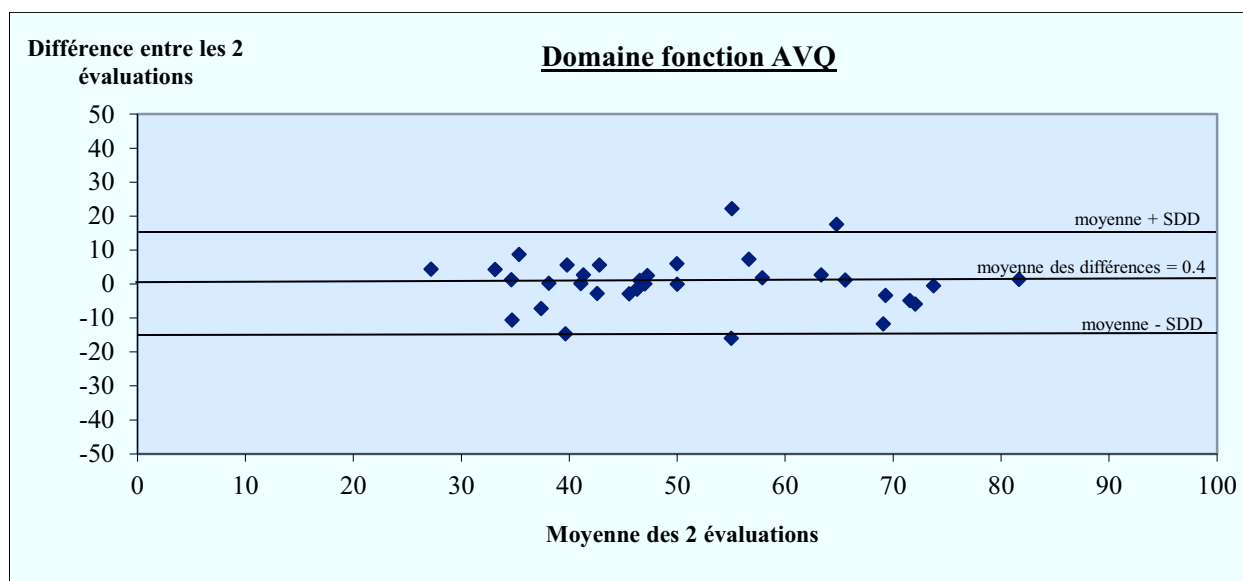
▪ **Figure 1a.**



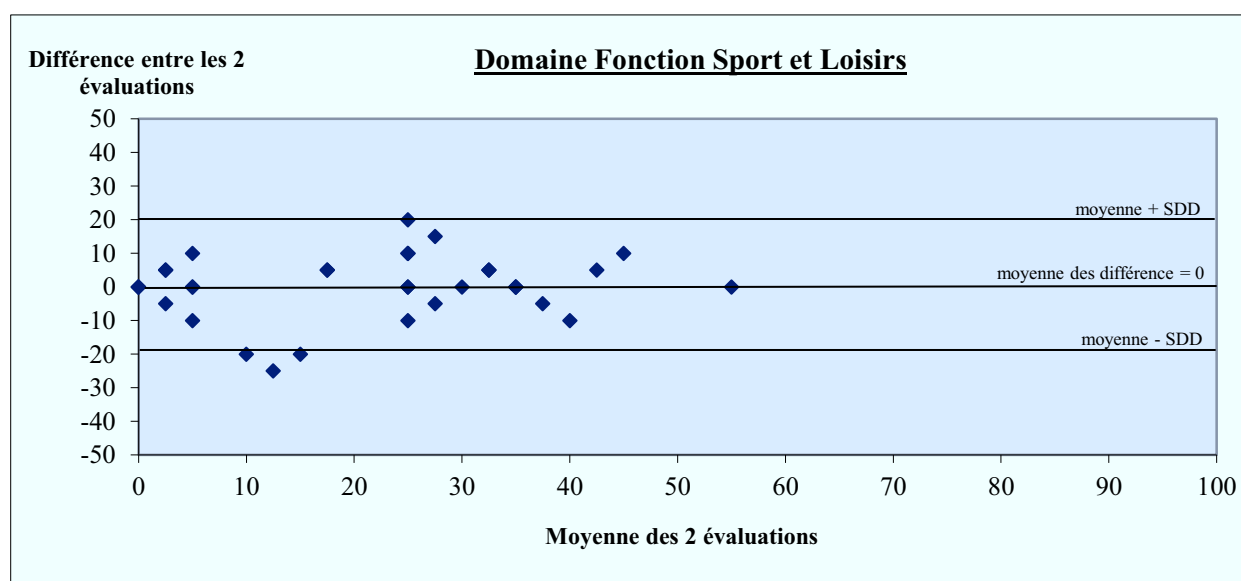
▪ **Figure 1b.**



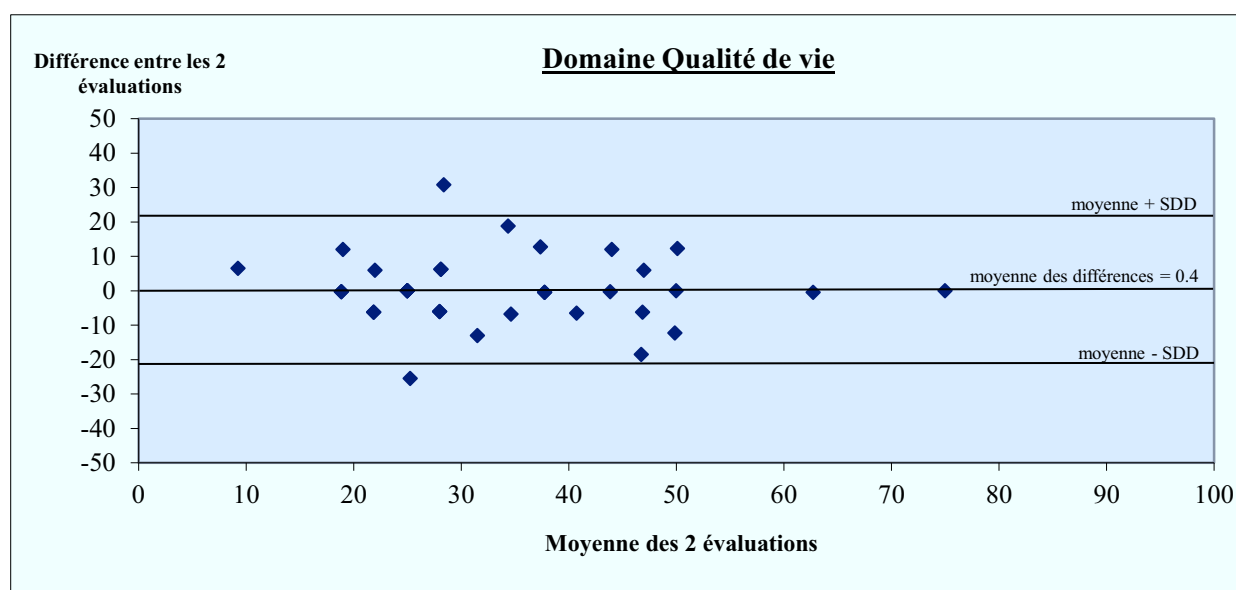
▪ Figure 1c.



▪ Figure 1d.



▪ **Figure 1e.**



La sensibilité au changement était élevée pour tous les domaines, avec une RMS allant de 0.89 à 1.93 et une taille de l'effet comprise entre 1.31 et 2.8 (tableau 5 et figure 2, page suivante). Le domaine *fonction SL* apparaissait comme le moins sensible au changement.

Domaines du KOOS	Taille de l'effet	RMS
<i>Douleur</i>	2.59	1.85
<i>Symptômes</i>	1.63	1.45
<i>Fonction AVQ</i>	2.52	1.8
<i>Fonction SL</i>	1.31	0.89
<i>QdV</i>	2.8	1.93

Tableau 5.Sensibilité au changement du KOOS

Les patients ont été évalués avant et trois après prothèse totale de genou.

RMS : Réponse moyenne standardisée

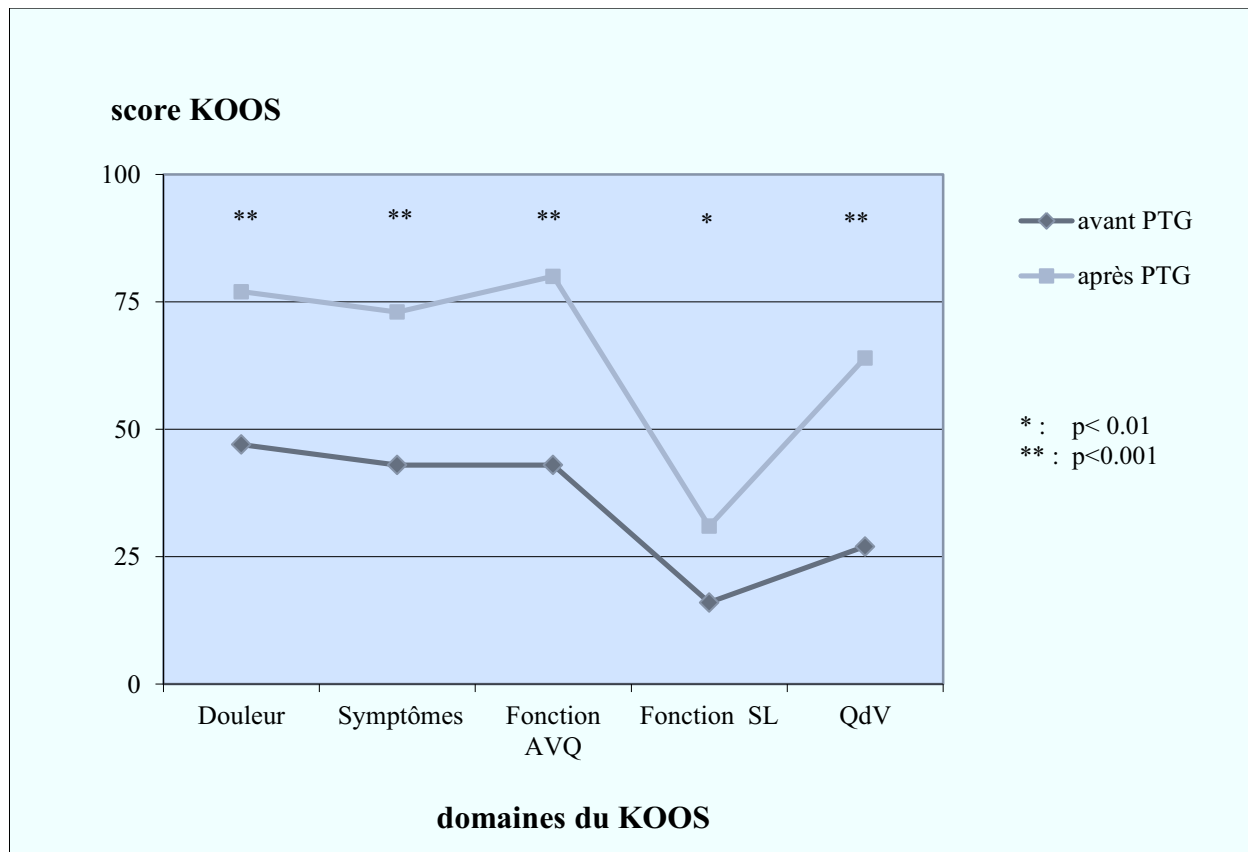


Figure 2. Sensibilité au changement du KOOS (n = 30).

scores des domaines du KOOS avant et 3 mois après prothèse totale de genou (PTG).

(0 = retentissement maximal ; 100 = retentissement nul)

2.5 Discussion

Dans cette étude, l'adaptation transculturelle du KOOS a été réalisée en français. Les propriétés psychométriques de la version traduite ont été ensuite évaluées avec des résultats jugés satisfaisants. Ces résultats devront être confirmés par d'autres études évaluant des sous-groupes différents, avant d'être généralisés. La version française est maintenant disponible et peut être téléchargée sur internet (217).

Les propriétés psychométriques de la version française du KOOS étaient généralement similaires à celles de la version originale (29, 228, 229). Après adaptation, la version française semble être un instrument de mesure avec une bonne faisabilité, comme en témoigne le faible

pourcentage de réponses manquantes. Les scores moyens du domaine *qualité de vie* et surtout du domaine *fonction SL* étaient sensiblement plus faibles que les scores des autres domaines, comme précédemment rapporté (29, 228, 234). Ceci pourrait être expliqué par le fait que les patients souffrant de gonarthrose évitent de pratiquer certaines activités critiques dans leur vie quotidienne, comme sauter ou courir. Cependant, ce résultat pourrait aussi être dû à l'âge des patients (âge moyen de 70 et 71 ans respectivement pour le groupe *médecine* et *chirurgie*). De plus, le domaine *fonction SL* n'a pas pu être calculé chez certains patients en pré et postopératoire. Les commentaires de certains patients suggèrent qu'ils n'ont pas répondu à certaines questions de ce domaine car ils ne pratiquaient pas ce type d'activités dans leur vie quotidienne ou qu'ils n'avaient pas encore essayé de les reprendre lors de l'évaluation post-chirurgicale. Le nombre important de données manquantes pour ce domaine particulier avait déjà été noté dans une précédente population de patients avec gonarthrose, en préopératoire (228). Toutefois, il est important d'inclure ces questions relatives à des activités physiques plus difficiles sachant qu'elles sont particulièrement appropriées pour le suivi des patients après arthroplastie, même à plus long terme (5 ans) (218). Mise à part pour ce domaine particulier, le pourcentage de réponses manquantes est très faible, suggérant que la version française est un auto-questionnaire approprié pour ces patients. Un problème de format a été découvert car quelques patients ont oublié de remplir une page du questionnaire. Un reformatage a été réalisé (4 pages au lieu de 10) pour éviter ce type de perte de données.

La cohérence interne et la reproductibilité ont été jugées bonnes et comparables aux résultats observés dans les autres langages (29, 232, 234). De plus, d'après les représentations graphiques de Bland et Altman, la différence entre les mesures répétées ne dépendait pas de la moyenne des deux mesures. La comparaison des résultats entre les deux populations étudiées suggère que ce questionnaire a une bonne capacité discriminante pour des patients avec des gonarthroses de sévérité différente. Les résultats des corrélations montrent que la version

française du KOOS possède une validité de construit convergente et divergente satisfaisante. Comme attendu, les plus fortes corrélations ont été trouvées entre les domaines du KOOS et de l'AMIQUAL mesurant les mêmes concepts (douleur et fonction). Les corrélations entre les domaines fonction et douleur confirment que les capacités fonctionnelles sont clairement associées à la douleur chez les patients avec une gonarthrose, comme cela avait déjà été rapporté (231, 235). Les corrélations entre le domaine *fonction SL* du KOOS et les autres domaines de l'AMIQUAL étaient faibles, ce qui pourrait paraître étonnant pour les domaines *douleur* ou *activités physiques*. Cependant, cela avait déjà été observé entre ce domaine du KOOS et le SF-36 (dimensions *activité physique* et *douleur physique*) (231, 235). Les corrélations fortes à modérées entre le domaine *santé mentale* de l'AMIQUAL et les domaines du KOOS (sauf pour le domaine *fonction SL*) pourraient être considérées comme surprenantes. Ce résultat n'avait pas été observé entre le SF-36 et le KOOS (228). Toutefois, de fortes corrélations avaient été rapportées entre le domaine *santé mentale* de l'AMIQUAL et le WOMAC fonction ou le WOMAC douleur (179). Contrairement au SF-36, l'AMIQUAL est particulièrement adapté pour mesurer les altérations de qualité de vie liées spécifiquement à la gonarthrose ou à la coxarthrose. Ceci pourrait expliquer les différences observées concernant les corrélations entre d'une part, le SF-36 et d'autre part, l'AMIQUAL. L'interprétation de l'ensemble des corrélations confirme que le domaine *qualité de vie* du KOOS capture plus que la douleur ou l'impotence fonctionnelle. D'autres résultats n'étaient pas attendus mais a posteriori ne sont pas si surprenants. Une corrélation modérée a été observée entre le domaine *fonction AVQ* du KOOS et le domaine *activités sociales* de l'AMIQUAL. Le domaine *activités sociales* de l'AMIQUAL contient deux questions qui sont en rapport avec le retentissement fonctionnel dans la vie quotidienne, ce qui peut expliquer ce résultat. Le domaine *qualité de vie* du KOOS n'est pas été corrélé avec le domaine *soutien*

social de l'AMIQUAL, sachant que cette dimension sociale n'est pas prise en compte dans les quatre questions du domaine *qualité de vie* du KOOS.

La version française a démontré une bonne sensibilité au changement. Le domaine *fonction SL* semble être moins sensible au changement que les autres domaines. Certes, l'évaluation post-chirurgicale n'a été réalisée que 3 mois après la chirurgie du genou. Toutefois, il est peu probable que cela explique ce résultat car cette plus faible sensibilité au changement avait été aussi rapportée à 6 mois et 12 mois après chirurgie dans une autre étude (228). De plus, pour tous les domaines, les résultats en termes de RMS et de taille de l'effet observés dans cette étude sont comparables à ceux décrits à 6 et 12 mois, suggérant que le choix d'une évaluation post-chirurgicale à 3 mois ne compromettrait pas les résultats. Par ailleurs, il faut souligner que la prothèse totale de genou est habituellement un traitement très efficace et que la RMS et la taille de l'effet n'auraient pas été aussi satisfaisants si elles avaient été évaluées pour un traitement moins efficace.

3. ADAPTATION TRANSCULTURELLE ET VALIDATION DE LA VERSION FRANÇAISE DU HOOS DANS LA COXARTHROSE.

3.1 Résumé

Objectif. Traduire et adapter le *Hip disability and Osteoarthritis Outcome Score* (HOOS) en français et évaluer les propriétés psychométriques de la version française, en testant sa faisabilité, sa cohérence interne, sa validité de construit, sa reproductibilité et sa sensibilité au changement chez des patients atteints de coxarthrose.

Méthodes. La version française du HOOS a été développée selon les recommandations internationales afin obtenir une validité de contenu adéquate. Cette version a ensuite été évaluée parmi deux populations de patients avec une coxarthrose symptomatique, une sans indication de prothèse totale de hanche (PTH) (groupe *médecine*) et l'autre en attente de pose de PTH (groupe *chirurgie*). Les propriétés psychométriques suivantes ont été évaluées : *faisabilité* par le pourcentage de réponses et la recherche d'effet plafond ou plancher ; la *cohérence interne* par le calcul de coefficient alpha de Cronbach, la *validité de construit* par la recherche de corrélations avec l'indice de Lequesne et l'EVA douleur (coefficient de corrélation de Spearman) ; la *reproductibilité* par le calcul du coefficient de corrélation intra-classe (ICC) et par la méthode graphique de Bland et Altman ; la *sensibilité au changement* en utilisant les données obtenues avant et un mois après viscosupplémentation (groupe *médecine*) et avant et trois mois après la pose de PTH (groupe *chirurgie*), par le calcul de la taille de l'effet et de la réponse moyenne standardisée (RMS).

Résultats.

88 patients ont été inclus au total ; 58 dans le groupe médecine (âge moyen = 61.8 ± 9 ans, extrêmes 42-81 ans, 70% femmes) et 30 dans le groupe chirurgie (âge moyen = 67.5 ± 9 ans, extrêmes 50-81 ans, 68% femmes). Le pourcentage de réponses était excellent (99%). Aucun

effet plafond ou plancher n'a été observé, mis à part un effet plafond avant la chirurgie pour le domaine *fonction sport et loisirs* (17.8% des patients avec le score le plus bas). La cohérence interne a été bonne pour 4 des 5 domaines du HOOS. Comme attendu, de fortes corrélations ont été observées entre tous les domaines du HOOS et l'indice de Lequesne ainsi qu'avec l'EVA douleur, ce qui témoigne d'une bonne validité de construit. La reproductibilité de mesure était bonne, avec des ICC supérieurs à 0.8 pour tous les domaines du HOOS. La sensibilité au changement pour la viscosupplémentation de hanche était bonne pour tous les domaines (taille de l'effet comprise entre 0.73 to 1.86 et RMS comprise entre 0.51 et 1.04) et était élevée pour la PTH (taille de l'effet entre 1.47 et 2.08 et RMS entre 1.97 et 3.24).

Conclusion Les propriétés psychométriques de la version française du HOOS sont bonnes, ce qui permet d'envisager son utilisation à grande échelle, pour évaluer des patients atteints de coxarthrose dans les essais cliniques, quelle que soit la sévérité de la coxarthrose.

3.2 Introduction

L'arthrose de hanche ou coxarthrose est l'une des principales causes de handicap dans le monde, avec un retentissement significatif sur la qualité de vie des patients. L'évaluation clinique des patients souffrant de coxarthrose nécessite d'évaluer la douleur, le retentissement fonctionnel et la qualité de vie en utilisant des questionnaires, soit spécifiques de la maladie, soit génériques (13, 213, 222, 226). Le questionnaire spécifique de l'arthrose le plus couramment utilisé dans le monde est le *Western Ontario and Mc Master Universities Index* (WOMAC) (27) qui prend en compte l'évaluation de la douleur, de la raideur et de l'impotence fonctionnelle.

Le *Hip disability and Osteoarthritis Outcome Score* (HOOS) est un auto-questionnaire dérivé du WOMAC, conçu pour les patients souffrant de coxarthrose ou d'une autre

pathologie de hanche (29-31). Ce questionnaire comprend cinq domaines : les domaines *douleur*, *symptômes*, *fonction dans les activités de la vie quotidienne (fonction AVQ)*, *fonction dans les activités de sports et de loisirs (fonction SL)* et *qualité de vie*. Le domaine *douleur* du HOOS contient les 5 questions du WOMAC douleur plus 5 autres questions. Le domaine *symptômes* contient les 2 questions du WOMAC raideur plus 3 autres questions relatives à la présence de bruit ou de craquement, à la difficulté pour écarter les jambes ou pour marcher à grandes enjambées. Le domaine *fonction AVQ* correspond aux 17 questions du WOMAC fonction. Les 2 derniers domaines, à savoir *fonction SL* (4 questions) et *qualité de vie* (4 questions sur la gêne, le manque de confiance, les modifications de la façon de vivre et la conscience du problème de hanche), sont deux nouveaux domaines qui ont pour objectif d'évaluer le retentissement de la coxarthrose sur des activités plus sollicitatrices et sur la qualité de vie. Pour chaque domaine, le score est normalisé sur 100 (100 correspondant à un retentissement nul et 0 à un retentissement maximal). Ces résultats peuvent ensuite être représentés sous la forme d'un graphique avec une courbe reliant les scores de chaque domaine. La validité, la reproductibilité et la sensibilité au changement de la version originale du HOOS ont été démontrées chez les patients avec une coxarthrose, ce qui permet d'envisager son utilisation comme critère d'évaluation dans cette pathologie (30, 31).

Du fait du nombre croissant d'études multicentriques internationales et de la nécessité d'obtenir des résultats significatifs pour les études épidémiologiques et/ou thérapeutiques, il est indispensable de réaliser l'adaptation et la validation transculturelle des instruments de mesure de la santé (22). Cette démarche spécifique nécessite non seulement la traduction mais parfois la transformation substantielle de certaines questions pour prendre en compte de manière exhaustive l'essence du concept original (18, 20). Idéalement, il faut ensuite que les propriétés psychométriques suivantes du questionnaire traduit soient vérifiées, comme définit par le filtre OMERACT : sa validité (pour refléter ce qu'il est censé mesurer),

sa reproductibilité, sa sensibilité au changement et sa faisabilité. Des versions validées du HOOS ont été publiées et sont disponibles en danois, hollandais (156), allemand, anglais, lituanien et suédois (217). Lors d'une analyse systématique récente de la littérature sur les questionnaires spécifiques de l'arthrose, Veenhof concluait que le HOOS était parmi les trois questionnaires avec les meilleurs résultats en termes de qualité psychométrique pour évaluer la douleur et l'impotence fonctionnelle dans la coxarthrose (67). Il est donc indispensable de poursuivre la traduction et l'adaptation transculturelle de ce questionnaire dans d'autres langages. L'objectif de cette étude était de réaliser la traduction française du questionnaire HOOS et de l'adapter en fonction des spécificités culturelles, puis d'évaluer ses propriétés psychométriques, en terme de faisabilité, de cohérence interne, de reproductibilité, de validité de construit et de sensibilité au changement chez des patients avec une coxarthrose symptomatique.

3.3 Méthodes

Dans un premier temps, le HOOS a été traduit et adapté en français. Dans un deuxième temps, les propriétés psychométriques de la version française ont été étudiées dans une étude prospective pour vérifier sa faisabilité, sa cohérence interne, sa validité de construit, sa reproductibilité et sa sensibilité au changement.

Traduction et adaptation transculturelle

La traduction et l'adaptation transculturelle depuis la version anglaise originale ont été réalisées selon les recommandations internationales (18, 20, 22).

Préparation. Le coordonnateur du projet (JFM) a contacté l'auteur du questionnaire (ER) pour obtenir la permission de l'utiliser et de le traduire, pour s'assurer qu'aucune autre

traduction française n'était en cours et pour lui proposer de participer à ce travail. Ils se sont ensuite rencontrés pour bien définir les notions comprises dans certaines questions, utiles pour les traducteurs.

Traduction. Trois personnes (deux rhumatologues, dont le coordonateur du projet, et un professeur d'anglais, d'origine française) ont traduit de manière indépendante la version anglaise du HOOS en français. Avant de commencer la traduction, tous les traducteurs ont été informés (en particulier celui qui n'était pas rhumatologue) sur la maladie cible, sur les concepts mesurés et divers autres renseignements qu'avait fournis l'auteur du questionnaire.

Réunion de concertation. Une version unique du HOOS en français a été obtenue après une réunion de concertation entre les trois traducteurs.

Rétro-Traduction. Cette nouvelle version a été traduite en anglais par une personne bilingue, d'origine anglaise, en aveugle de la version anglaise originale.

Revue de la rétro-traduction et harmonisation. Le coordonateur du projet a comparé la version anglaise ainsi obtenue à la version originale, à la recherche de différences ou de contradictions. Une réunion multidisciplinaire a été ensuite organisée avec les trois traducteurs, un autre rhumatologue (PO), un chirurgien orthopédique (YJ) et un rhumatologue-épidémiologiste spécialisé dans l'adaptation transculturelle (FG), pour réaliser une revue des traductions réalisées et ainsi harmoniser les résultats. Durant cette réunion, le comité vérifia si la traduction était compréhensible dans son ensemble et si les équivalences transculturelles entre les versions anglaise et française étaient respectées, ce qui permit d'aboutir, par consensus, à une nouvelle version en français.

Pré-test et revue. Cette version a été pré-testée chez 15 patients qui venaient consulter en Rhumatologie au centre hospitalier universitaire de Dijon pour coxarthrose, avec des caractéristiques comparables à celles des patients inclus dans la seconde partie de l'étude.

Les patients devaient remplir le questionnaire en présence d'un des rhumatologues associés au projet. On leur demandait s'ils comprenaient bien toutes les questions et si leur formulation pouvait poser problème. Les résultats du pré-test étaient revus par le coordonnateur du projet et si nécessaire, des corrections étaient apportées.

Validation de la version finale. Lors de cette étape finale, le coordonnateur du projet a vérifié que la version finale ne comportait pas d'erreur. Durant l'ensemble de la procédure, les participants devaient garder en mémoire que le questionnaire devait être compris par des non-spécialistes ou des individus avec un faible niveau d'éducation (équivalent au niveau d'un enfant de 10 à 12 ans). A tout moment, le coordonnateur du projet avait la possibilité de contacter l'auteur pour discuter de problèmes relatifs au questionnaire.

Propriétés psychométriques de la version française du HOOS

Patients.

Deux populations de patients atteints de coxarthrose ont été évaluées dans cette étude prospective bi-centrique. Le groupe *médecine* comprenait des patients qui consultaient pour une coxarthrose symptomatique dans le service de Rhumatologie de l'hôpital universitaire de Dijon. Le groupe *chirurgie* était constitué de patients qui consultaient dans le service de Chirurgie orthopédique des hôpitaux universitaires de Marseille pour la pose d'une prothèse de hanche (PTH). Les critères d'inclusion étaient un âge supérieur à 40 ans, une coxarthrose primitive répondant aux critères de l'*American College of Rheumatology* (236) et la capacité à comprendre et à remplir seul les auto-questionnaires. Chez les patients évalués pour la sensibilité au changement dans le groupe *médecine*, un critère supplémentaire était l'indication de viscosupplémentation de hanche par acide hyaluronique selon le rhumatologue traitant. Chez les patients évalués pour la sensibilité au changement dans le groupe *chirurgie*, l'indication de PTH devait être retenue par leur chirurgien.

Les critères d'exclusion étaient la présence d'une pathologie rhumatismale invalidante comme par exemple une arthrose symptomatique d'une autre articulation des membres inférieurs, une coxarthrose secondaire ou inflammatoire en poussée, des lombalgies chroniques ou la notion d'une infiltration intra-articulaire de corticoïdes durant les trois mois précédents. Chez les patients inclus dans l'étude de la reproductibilité, un critère d'exclusion supplémentaire était la modification du traitement de la coxarthrose dans les deux semaines à venir. Chez les patients évalués pour la sensibilité au changement dans le groupe *médecine*, un critère d'exclusion supplémentaire était la modification du traitement de l'arthrose au cours du prochain mois, hormis la viscosupplémentation.

Questionnaires

Groupe médecine

Durant la première évaluation, les patients devaient remplir la version française du HOOS, l'indice de Lequesne et évaluer leur douleur sur une échelle visuelle analogique (EVA douleur). Cinq possibilités de réponses (sur une échelle de Likert) étaient proposées pour chaque question du HOOS, cotées entre 0 et 4. Pour chaque domaine du HOOS, les scores étaient normalisés sur une échelle entre 0 et 100 (0 = retentissement maximal, 100 = retentissement nul). Les données manquantes étaient prises en compte selon les recommandations formulées dans la version originale du HOOS (29, 222) : quand plus de deux réponses d'un domaine étaient manquantes, le score total n'était pas calculé. Dans les autres situations, les données manquantes étaient remplacées par le score moyen obtenu dans le même domaine par le patient. L'indice de Lequesne (60, 111) est un questionnaire spécifique de l'arthrose comprenant 10 questions qui permet d'évaluer la douleur et l'impotence fonctionnelle liée à la coxarthrose. Ce questionnaire, dont le score va de 0 à 24 (un score faible témoignant de répercussions modérées) est couramment utilisé comme critère

d'évaluation dans l'arthrose des membres inférieurs. L'EVA douleur est un outil largement utilisé et validé pour mesurer l'intensité de la douleur (13). Les patients évalués pour l'étude de la reproductibilité de mesure devaient remplir un second questionnaire HOOS 15 jours plus tard et devaient le renvoyer par courrier, dans une enveloppe préaffranchie. Un intervalle de 15 jours a été jugé suffisamment long pour s'assurer de l'absence de mémorisation des réponses et suffisamment court pour éviter une modification dans l'intervalle de la douleur et de l'impotence fonctionnelle dues à la coxarthrose. Les patients évalués pour la sensibilité au changement bénéficiaient d'une viscosupplémentation de hanche par acide hyaluronique, sous contrôle échographique. La procédure était réalisée par le même praticien (PO). La nature et le poids moléculaire de l'acide hyaluronique variaient car les patients étaient adressés avec la prescription du produit déjà effectuée par leur rhumatologue traitant. Ces patients devaient remplir un deuxième questionnaire HOOS un mois plus tard et le renvoyer dans une enveloppe préaffranchie.

Groupe chirurgie

Les patients devaient remplir le questionnaire HOOS avant la pose de prothèse totale de hanche et trois mois après l'opération, lors d'une consultation de suivi.

Analyse statistique

Faisabilité

La faisabilité a été évaluée en utilisant le pourcentage de données manquantes et en recherchant un effet plafond et un effet plancher dans les deux groupes. La présence d'un effet plafond ou plancher peut influencer la reproductibilité, la validité et la sensibilité au changement d'un instrument de mesure. Dans cette étude, les effets plafond et plancher ont été considérés présents si 15% des patients interrogés atteignaient le score le plus bas (retentissement maximal) ou le plus élevé (retentissement nul).

Cohérence interne

La cohérence interne permet d'évaluer les interrelations entre les questions d'un même domaine et leur degré d'homogénéité. La cohérence interne a été évaluée en utilisant le coefficient de Cronbach (55), avec les intervalles de confiance à 95%. Des divergences existent dans la littérature pour le seuil du coefficient de Cronbach considéré comme satisfaisant (81, 237). Dans cette étude, un coefficient > 0.7 a été jugé satisfaisant (81).

Validité de construit

Dans le groupe *médecine*, la validité de construit a été évaluée en corrélant les résultats des cinq domaines du HOOS avec les résultats de l'indice de Lequesne et de l'EVA douleur. Comme cela est recommandé dans la littérature (238), des coefficients de Pearson > 0.50 , entre 0.35 et 0.50 et < 0.35 ont été considérés respectivement comme des corrélations forte, modérée et faible. Les hypothèses suivantes ont été formulées a priori : tous les domaines du HOOS devraient être corrélés fortement avec l'indice de Lequesne et l'EVA douleur. Les plus fortes corrélations devraient être retrouvées entre les domaines évaluant le même concept (EVA douleur et domaine *douleur* du HOOS, indice de Lequesne et domaine *fonction AVQ* du HOOS).

Reproductibilité

La reproductibilité des différents domaines du HOOS a été évaluée en utilisant le coefficient de corrélation intra-classe (ICC), avec les intervalles de confiance à 95%. Un ICC supérieur à 0.80 est habituellement considéré comme le témoin d'une bonne reproductibilité de mesure (238, 239). Par ailleurs, la méthode graphique de Bland et Altman (56), dans laquelle la différence entre la première et la deuxième évaluation est représentée (en ordonnée) en fonction de la moyenne de ces deux évaluations (en abscisse), a été obtenue pour chaque domaine du HOOS. La plus petite différence perceptible (SDD, *smallest detectable*

difference en anglais) qui correspond à 1.96 fois l'écart-type des différences entre les mesures répétées a ainsi pu être calculée pour chaque domaine. Le SDD correspond à la plus petite différence qui peut être distinguée de l'erreur de mesure (56, 57).

Sensibilité au changement

La sensibilité au changement a été évaluée en comparant les résultats avant et 1 mois après viscosupplémentation dans le groupe *médecine* et avant et 3 mois après PTH dans le groupe *chirurgie*. La réponse moyenne standardisée (RMS), qui correspond à la différence moyenne entre le score avant et après traitement divisée par l'écart type de cette différence et la taille de l'effet qui correspond à la différence moyenne entre le score avant et après traitement divisée par l'écart-type du score avant traitement ont ainsi été calculées pour chaque domaine.

Le logiciel SPSS (*Statistical Package for the Social Sciences* (SPSS) version 14.0) a été utilisé pour l'analyse statistique. Le degré de significativité était défini pour un $p < 0.05$.

3.4 Résultats

Traduction et adaptation transculturelle

Des différences minimales dans la structure des questions ont été mises en évidence entre les trois traductions en français, ainsi qu'entre la version anglaise originale et celle rétro-traduite en anglais. Avant d'obtenir un consensus, chaque mot ou expression a été longuement discuté par le comité pour conserver leur sens et pour que le questionnaire soit bien compréhensible par tous les patients. Lors du pré-test, une grande majorité des patients a trouvé que le questionnaire était facile à comprendre et à remplir. Dans quelques cas, il a été nécessaire de préciser que les questions correspondaient à leur état au cours de la dernière

semaine. Trois patients ont trouvé que le questionnaire contenait trop de questions. Aucune modification n'a été apportée au questionnaire après le pré-test.

Caractéristiques des patients

Au total, 88 patients ont été inclus : 58 dans le groupe *médecine* (âge moyen = 61.8 ± 9 ans, extrêmes entre 42 et 81 ans, 70% femmes) and 30 dans le groupe *chirurgie* (âge moyen = 67.5 ± 9 ans, extrêmes entre 50 et 81 ans, 68% femmes). La répartition des patients pour l'évaluation des propriétés psychométriques du questionnaire s'est fait comme suit :

- La faisabilité a été étudiée au moyen du premier questionnaire HOOS, rempli par l'ensemble des 88 patients de l'étude.
- La cohérence interne a été évaluée en utilisant le premier HOOS (88 patients).
- La validité de construit a été évaluée chez les 32 patients du groupe *médecine* qui avaient rempli à la fois l'indice de Lequesne et l'EVA douleur, en plus du HOOS.
- La reproductibilité a été évaluée chez les patients ayant répondu aux deux questionnaires à 15 jours d'intervalle dans le groupe *médecine* (43 patients sur 58). Les patients restants (15 sur 58) avaient bénéficié d'une viscosupplémentation de hanche par acide hyaluronique à la première visite, donc avaient été exclus de l'étude de la reproductibilité.
- La sensibilité au changement incluait les 30 patients du groupe *chirurgie* et 21 patients du groupe *médecine* ayant bénéficié d'une viscosupplémentation (les 15 patients qui n'avaient pas participé à la reproductibilité et 6 autres patients après avoir participé à l'étude de la reproductibilité). La grande majorité des patients (28/30 dans le groupe *chirurgie* et 21/21 dans le groupe *médecine*) ont complété puis renvoyé leur second questionnaire.

Propriétés psychométriques

a. Faisabilité

Dans le groupe *médecine* et le groupe *chirurgie* avant la PTH (88 patients), peu de réponses étaient manquantes (respectivement 0.9% and 1.32%) Le score total pour chaque domaine a pu être calculé pour tous les patients. Aucun effet plancher ou plafond n'a été observé dans le groupe *médecine*. Dans le groupe *chirurgie*, un effet plafond a été observé avant la chirurgie (17.8% des patients avec un retentissement maximal pour le domaine *fonction SL*). Cet effet plafond disparaissait après la pose de PTH (0%).

b. Cohérence interne

Le tableau 1 présente les résultats pour les 88 patients (groupes *médecine* et *chirurgie*).

Domaines du HOOS (nombre de questions)	Coefficient alpha de Cronbach (IC 95%)
<i>Douleur (10)</i>	0.86 (0.79-0.91)
<i>Symptômes (5)</i>	0.66 (0.53-0.76)
<i>Fonction AVQ(17)</i>	0.94 (0.91-0.96)
<i>Fonction SL (4)</i>	0.85 (0.78-0.90)
<i>Qualité de vie (4)</i>	0.82 (0.74-0.88)

Tableau 1. Cohérence interne des différents domaines du HOOS (n = 88)

IC 95% : intervalle de confiance à 95%

Le coefficient alpha de Cronbach allait de 0.82 jusqu'à 0.94, témoin d'une bonne homogénéité, excepté pour le domaine *symptômes* pour lequel le Cronbach était en dessous de 0.7 (0.66, 95% CI = 0.53-0.76).

La cohérence interne du domaine raideur du WOMAC (qui est compris dans le domaine *symptômes* du HOOS) était faible (0.50). Le coefficient de Cronbach du domaine douleur du WOMAC (0.81) était équivalent à celui du domaine *douleur* du HOOS (0.86).

c. Validité de construit

Le tableau 2 présente les corrélations entre les domaines du HOOS, l'indice de Lequesne et l'EVA douleur.

HOOS	<i>Douleur</i>	<i>Symptômes</i>	<i>Fonction AVQ</i>	<i>Fonction SL</i>	<i>QdV</i>
<i>EVA douleur</i>	-0.67 (p < 0.001)	-0.49 (p = 0.004)	-0.71 (p < 0.001)	-0.58 (p < 0.001)	-0.55 (p = 0.001)
<i>Indice de Lequesne</i>	-0.68 (p < 0.001)	-0.54 (p = 0.002)	-0.69 (p < 0.001)	-0.82 (p < 0.001)	-0.51 (p = 0.003)

Tableau 2. Validité de construit du HOOS.
Corrélations entre les domaines de la version française du HOOS, l'EVA douleur et l'indice de Lequesne pour la coxarthrose
(Coefficient de Pearson), n = 32.

De fortes corrélations ont été observées entre le HOOS et l'indice de Lequesne ainsi qu'avec l'EVA douleur. Comme attendu, les plus fortes corrélations ont été obtenues entre les domaines mesurant les mêmes concepts comme la fonction (indice de Lequesne et le domaine *fonction AVQ* (r = -0.69) ou le domaine *fonction SL* (r = -0.82)) ou la douleur (EVA douleur et domaine *douleur* du HOOS (r = -0.67)). La corrélation entre l'EVA douleur et le domaine *fonction AVQ* était aussi forte (r = -0.71).

d. Reproductibilité

Le tableau 3 présente les résultats en termes de reproductibilité test-retest.

Domaines du HOOS	1^{ère} évaluation <i>moyenne (DS)</i>	2^{ème} évaluation <i>moyenne (DS)</i>	ICC (IC 95%)	SDD
<i>Douleur</i>	55.3 (14.2)	52.0 (13.7)	0.83 (0.70-0.90)	15.1
<i>Symptômes</i>	54.1 (14.9)	52.9 (14.5)	0.84 (0.72-0.91)	10.5
<i>Fonction AVQ</i>	53.9 (16.3)	53.0 (15.9)	0.86 (0.76-0.92)	9.6
<i>Fonction SL</i>	34.6 (20.4)	33.5 (20.5)	0.89 (0.80-0.94)	15.5
<i>QdV</i>	37.5 (19.5)	38.2 (18.3)	0.86 (0.76-0.92)	16.2

Tableau 3. Scores moyens du HOOS réalisés à 15 jours d'intervalle et leur reproductibilité de mesure en termes d'ICC (n = 43)

SDD : plus petite difference perceptible (*smallest detectable difference*)

ICC : coefficient de corrélation intra-classe

DS : déviation standard

IC 95% : intervalle de confiance à 95%

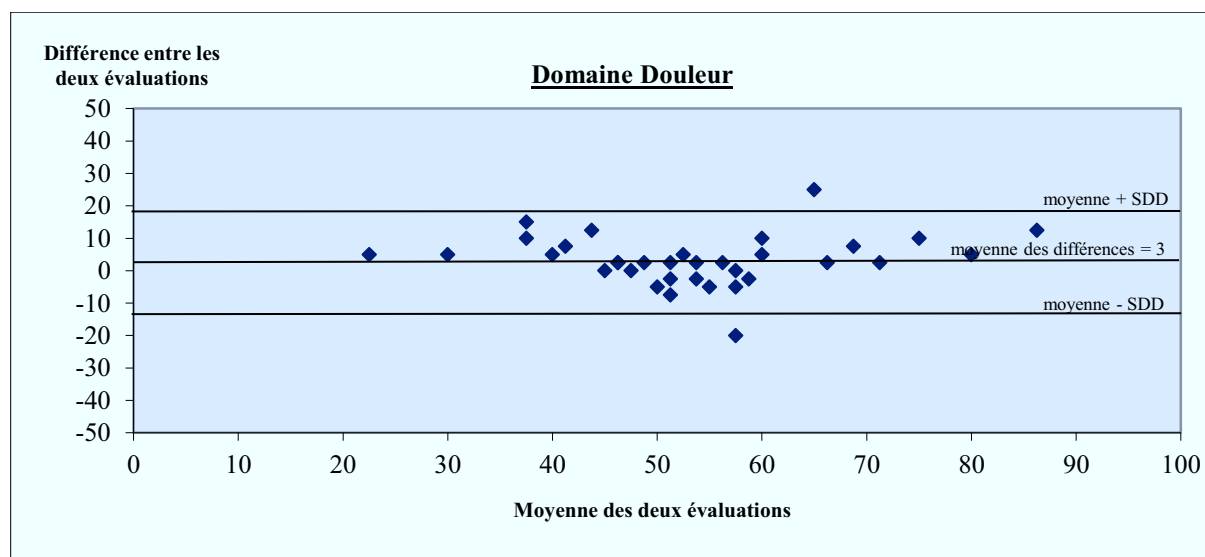
Pour tous les domaines du HOOS, les ICCs étaient bons, allant de 0.83 (domaine *douleur*) à 0.89 (domaine *fonction SL*). La reproductibilité était similaire entre le score douleur du WOMAC (ICC = 0.81) et le score du domaine *douleur* du HOOS (ICC = 0.83) et entre le domaine raideur du WOMAC (ICC = 0.82) et le domaine *symptômes* du HOOS (ICC = 0.84). Le SDD était compris entre 9.6 (domaine *fonction AVQ*) et 16.2 (domaine *qualité de vie*).

La méthode graphique de Bland et Altman est présentée pour chaque domaine dans la figure 1.

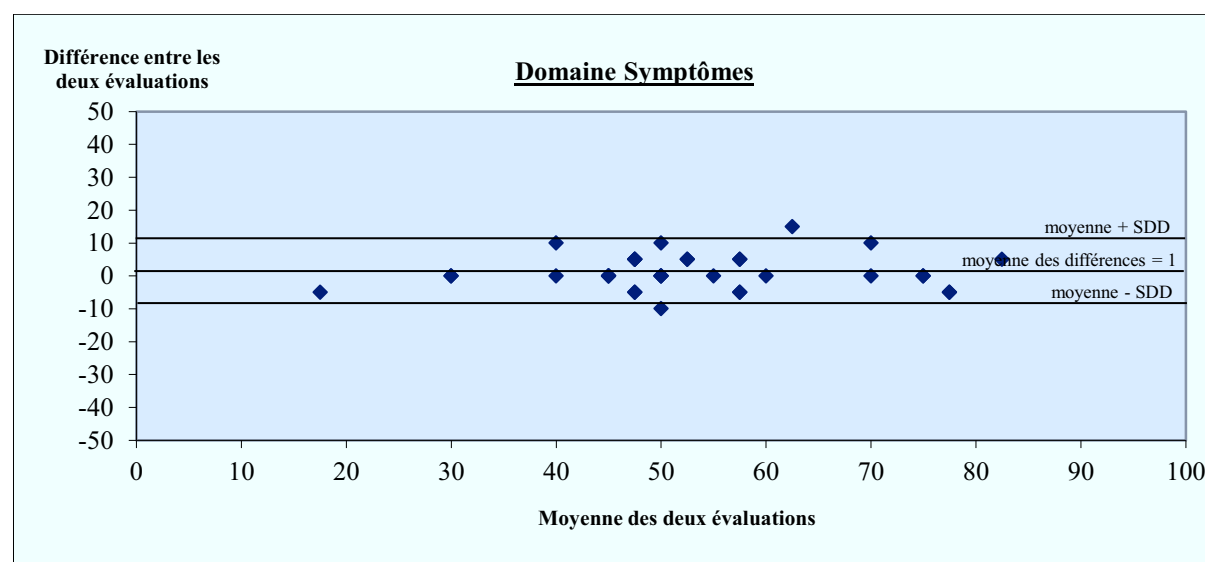
Figure 1.

Reproductibilité de la version française du HOOS (méthode graphique de Bland et Altman). Les deux évaluations successives ont été réalisées à 15 jours d'intervalle.

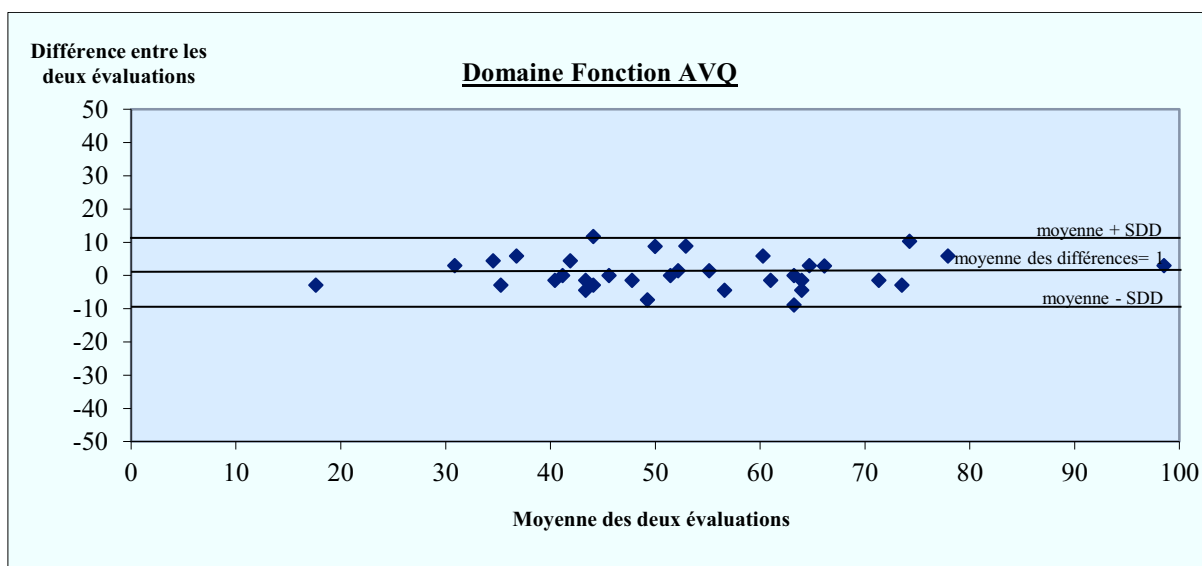
▪ **Figure 1a.**



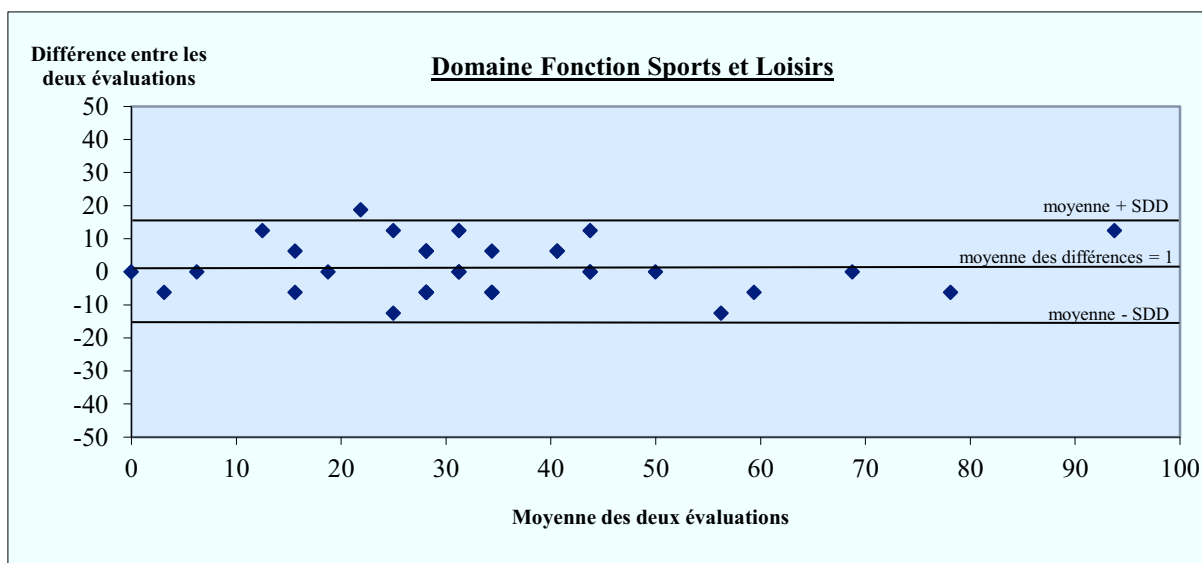
▪ **Figure 1b.**



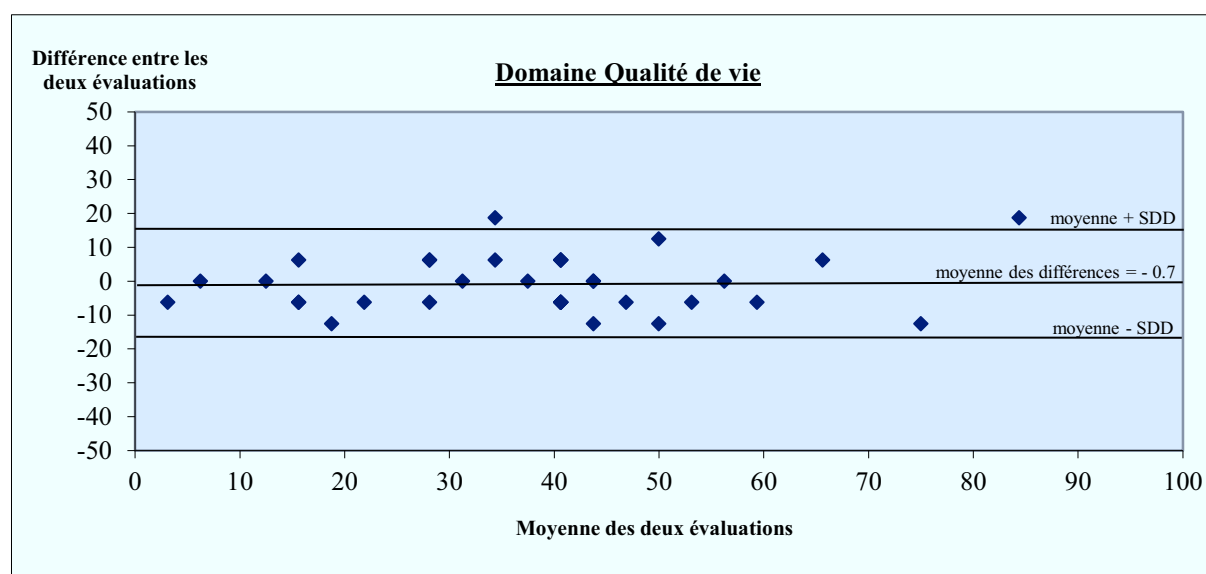
▪ **Figure 1c.**



▪ **Figure 1d.**



▪ **Figure 1e.**



La différence entre les deux mesures répétées du HOOS reste ainsi comprise entre les bornes de l'intervalle de confiance dans la grande majorité des cas et ne dépend pas du score moyen des deux mesures.

e. Sensibilité au changement

Les résultats pour la sensibilité au changement (PTH et viscosupplémentation de hanche) sont présentés dans la figure 2 et le tableau 4 des pages suivantes. Tous les scores des domaines du HOOS s'amélioraient significativement après la chirurgie ($p < 0.001$) et la viscosupplémentation ($p < 0.01$). Dans le groupe *chirurgie*, la sensibilité au changement était élevée pour tous les domaines, avec une RMS comprise entre 1.97 à 3.24 et une taille de l'effet comprise entre 1.47 à 2.08. Le domaine *qualité de vie* était moins sensible au changement que les autres domaines (figure 2a). Dans le groupe *médecine*, la sensibilité au changement était bonne après viscosupplémentation, avec une RMS comprise entre 0.51 à 1.04 et une taille de l'effet comprise entre 0.73 à 1.86. L'amélioration pour le domaine *fonction SL* était moins importante que pour les autres domaines (figure 2b).

Figure 2. Sensibilité au changement de la version française du HOOS

(0 = retentissement maximal ; 100 = retentissement nul)

Figure 2a. Domaines du HOOS avant et 3 mois après prothèse totale de hanche (PTH)

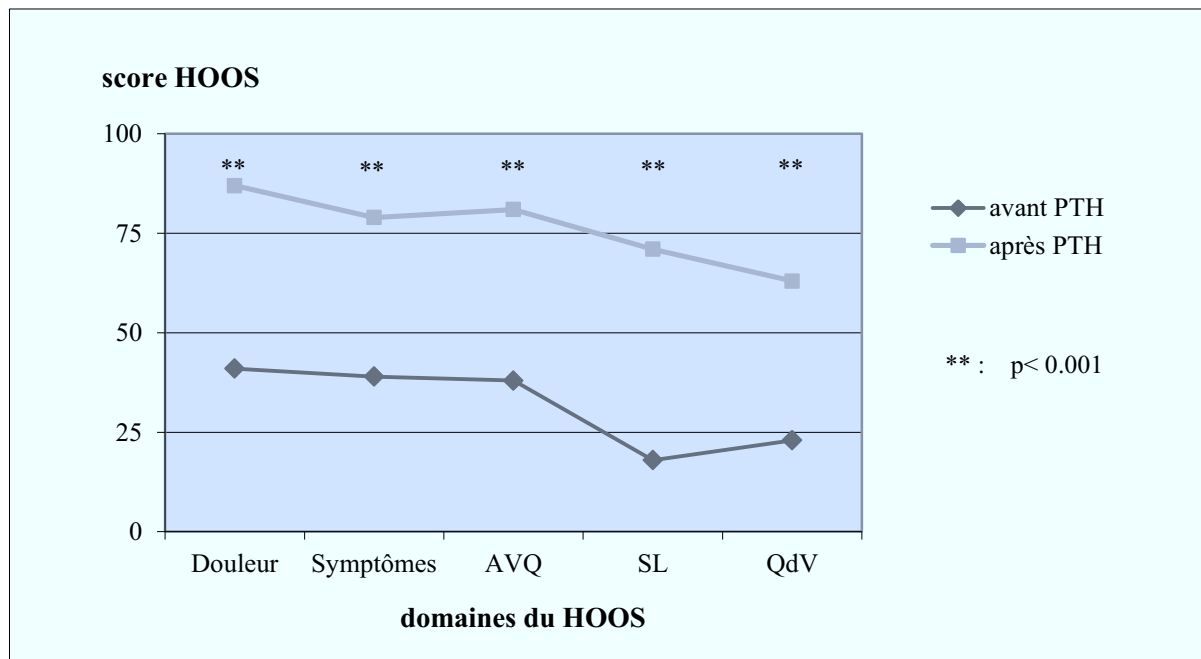
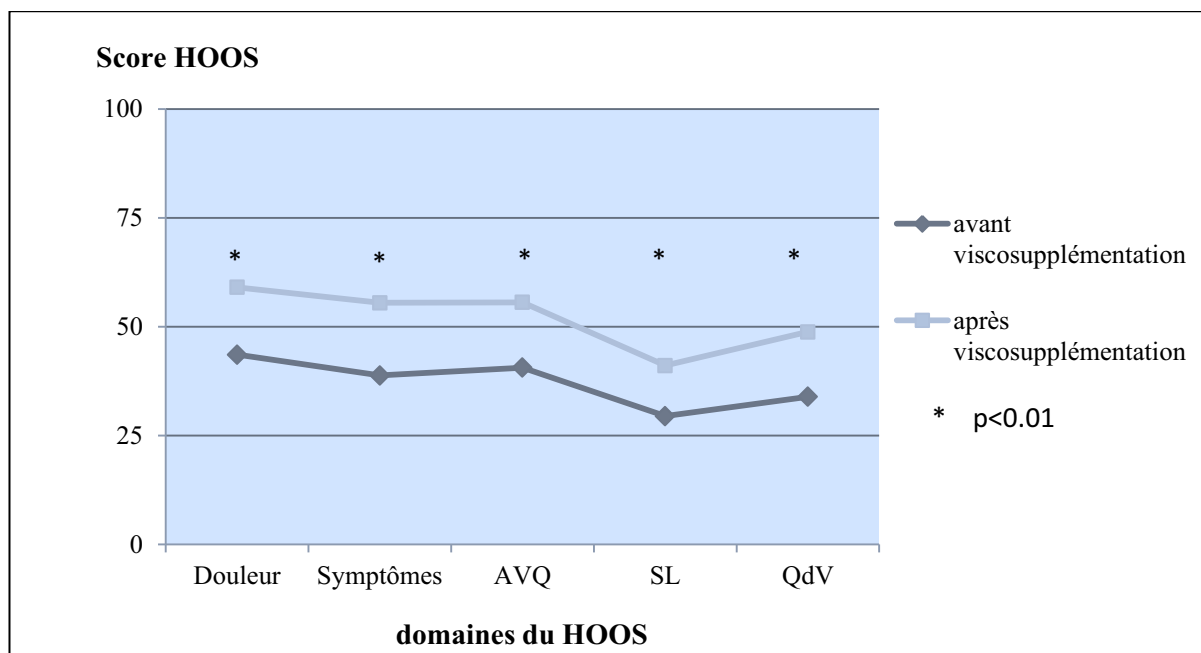


Figure 2b. Domaines du HOOS avant et 1 mois après viscosupplémentation de hanche



Dans les deux groupes, la sensibilité au changement était significativement plus importante pour le domaine *douleur* du HOOS (taille de l'effet égal à 1.86 pour la viscosupplémentation et égal à 3.24 pour la PTH) que pour le WOMAC douleur (taille de l'effet égal à 1.25 pour la viscosupplémentation et égal à 2.51 pour la PTH). Par ailleurs, la sensibilité au changement était plus importante pour le domaine *symptômes* du HOOS que pour le WOMAC raideur dans les deux groupes, mais sans atteindre le seuil de la significativité.

Domaines du HOOS	Prothèse totale de hanche n = 30				Viscosupplémentation de hanche n = 21			
	moyenne avant (DS)	moyenne après (DS)	Taille de l'effet	RMS	moyenne avant (DS)	moyenne après (DS)	Taille de l'effet	RMS
<i>Douleur</i>	41.7 (13.8)	86.5 (16.1)	3.24	1.85	43.6 (15.7)	59.0 (17.5)	1.86	0.99
<i>Symptômes</i>	41.0 (17.0)	77.3 (14.2)	2.14	1.54	38.8 (16.0)	55.5 (19.7)	1.22	1.04
<i>Fonction AVQ</i>	38.7 (14.7)	81.8 (13.5)	2.83	2.08	40.6 (16.2)	55.6 (18.0)	1.47	0.92
<i>Fonction SL</i>	18.3 (14.6)	64.0 (19.7)	3.11	1.90	29.5 (22.8)	41.1 (27.2)	0.73	0.51
<i>QdV</i>	21.4 (20.7)	61.4 (21.2)	1.97	1.47	33.9 (20.4)	48.8 (19.0)	1.00	0.73

Tableau 4.Sensibilité au changement de la version française du HOOS

Les 30 patients du groupe *chirurgie* ont été évalués avant et trois après prothèse totale de hanche. Les 21 patients du groupe *médecine* ont été évalués avant et 1 mois après la viscosupplémentation de hanche.

RMS : réponse moyenne standardisée

DS : déviation standard

3.5 Discussion

Dans cette étude, l'adaptation transculturelle de la version anglaise du HOOS a été réalisée en français. Les propriétés psychométriques de la version française ont ensuite été évaluées, avec des résultats satisfaisants et similaires à ceux de la version originale. La version française du HOOS n'est pas soumise au droit d'auteur et peut être téléchargée gratuitement sur le site du KOOS (www.KOOS.nu) (217). Tous les questionnaires ont été remplis et le faible pourcentage de données manquantes témoigne de la bonne acceptabilité et de la bonne faisabilité de la version française du HOOS. Les résultats de la cohérence interne sont excellents pour 4 des 5 domaines et comparables à ceux observés dans les autres langues (31, 156). Comme dans les précédentes études de validation du HOOS, le coefficient de Cronbach le plus élevé a été retrouvé pour le domaine *fonction AVQ*. Bien qu'acceptable, la cohérence interne pour le domaine *symptômes* était nettement inférieure (0.66) à celle observée avec les autres domaines (> 0.82). Ce résultat inattendu pourrait indiquer une moins bonne corrélation inter-items pour ce domaine. Contrairement au domaine *douleur*, le domaine *symptômes* pourrait capturer plus qu'un seul concept (bruits provenant de la hanche, difficultés ou sensation de raideur lors de son utilisation) et de fait, cette composante multidimensionnelle nécessite une interprétation plus prudente de la cohérence interne pour ce domaine particulier. Toutefois, ce résultat n'avait pas été rapporté jusqu'à présent (le coefficient de Cronbach le plus bas avait été retrouvé pour le domaine *qualité de vie* dans la version suédoise (30) et pour le domaine *douleur* dans la version hollandaise (156)). Une autre explication pourrait être un problème de traduction et d'adaptation transculturelle en français. Cependant, un tel résultat n'avait pas été observé dans une étude précédente évaluant les propriétés psychométriques de la version française du KOOS dans laquelle la cohérence interne du domaine *symptômes* était bonne (240), alors que plusieurs questions sont

communes ou très proches entre ces deux questionnaires. Finalement, l'explication pourrait être des variations d'échantillonnage sachant que la borne supérieure de l'intervalle de confiance du domaine du HOOS comprend 0.7. Quelle que soit l'explication retenue, il n'est pas actuellement possible d'affirmer que la version française du domaine *symptômes* est suffisamment homogène. D'autres études sont nécessaires pour répondre à cette question. En attendant, ce domaine devrait être utilisé plutôt au niveau du groupe plutôt qu'au niveau individuel qui nécessite un niveau plus élevé de cohérence interne (241).

Les scores moyens de deux domaines qui ne sont pas inclus dans le WOMAC (*fonction SL* et *qualité de vie*) étaient nettement plus bas que ceux des autres domaines du HOOS, en concordance avec les études de validation précédentes (31, 156). Ceci pourrait être expliqué par l'âge des patients de l'étude (âge moyen = 68 ans). La reproductibilité était excellente ($ICC > 80$) pour tous les domaines du HOOS, comme dans les études précédentes. La méthode graphique de Bland et Altman a permis de vérifier que les différences entre les mesures répétées n'étaient pas liées à la moyenne de ces mesures.

La validité de construit était satisfaisante : tous les domaines du HOOS étaient fortement corrélés avec l'indice de Lequesne et l'EVA douleur, avec les plus fortes corrélations retrouvées entre les domaines mesurant des concepts similaires, comme cela était attendu.

La sensibilité au changement est une propriété cruciale d'un instrument de mesure car une sensibilité au changement élevée permet de réduire le nombre de sujets nécessaire pour démontrer une différence significative entre des groupes. Dans cette étude, la sensibilité au changement du HOOS était bonne dans les deux échantillons étudiés. Ce résultat était attendu dans le groupe *chirurgie* car la PTH est un traitement très efficace dans la coxarthrose et car cette sensibilité au changement du HOOS pour la PTH avait déjà été considérée comme excellente dans d'autres populations de coxarthroses (31, 156). A l'inverse, d'après nos

connaissances, la sensibilité au changement du HOOS après viscosupplémentation par acide hyaluronique n'avait jamais encore été évaluée. Comme ce traitement est moins efficace que la PTH, et qu'il n'a pas démontré de différence significative avec l'injection de placebo (242, 243), une taille de l'effet ou une RMS faibles étaient plutôt attendus, alors que les résultats de la sensibilité au changement du HOOS pour ce traitement sont bons. Cependant, la sensibilité au changement pour la viscosupplémentation sur des séries de cas reste toujours difficile à interpréter en raison du faible nombre de patients, de l'absence de population contrôle, de l'effet placebo, des co-interventions potentielles et de l'absence d'aveugle pour les investigateurs (244). Une partie des modifications observées pourrait être due à un effet placebo (245), particulièrement important dans l'arthrose, surtout pour les traitements injectables. De plus, dans le groupe *médecine*, la deuxième évaluation réalisée un mois après l'injection d'acide hyaluronique pourrait expliquer un important effet placebo. Cela a été observé précédemment dans une étude (246) dans laquelle 25% des patients dans le groupe placebo étaient considérés comme répondeurs au bout d'un mois mais seulement 4.8% à 3 mois. De plus, les différences entre les groupes *chirurgie* et *médecine* concernant la prise d'analgésiques ou d'AINS limitent la comparaison des résultats pour la sensibilité au changement entre la PTH et la viscosupplémentation dans notre étude. En tout état de cause, l'objectif de cette étude n'était pas d'étudier l'efficacité de l'injection d'acide hyaluronique dans la coxarthrose mais d'évaluer les propriétés psychométriques du questionnaire HOOS. Concernant cet objectif, les résultats sont très satisfaisants ce qui permet d'envisager son utilisation dans les essais thérapeutiques. Par ailleurs, ces résultats suggèrent que la version française du HOOS peut détecter des effets moins importants que ceux observés avec la PTH et que ce questionnaire a une plus grande sensibilité au changement que le WOMAC chez ces patients. Le HOOS pourrait donc être intéressant pour évaluer des traitements habituels de la coxarthrose, dans différents groupes de patients.

Cette étude présente certaines limites : premièrement, les participants pourraient ne pas représenter l'ensemble du spectre des patients avec une coxarthrose car ils avaient été recrutés dans deux hôpitaux universitaires, avec des arthroses potentiellement plus sévères, en particulier dans le groupe *chirurgie*. Ceci pourrait expliquer en parti l'effet plafond observé avant la pose de PTH pour le domaine *fonction SL*, qui avait été développé initialement pour des patients plus jeunes et plus actifs. Cet effet plafond avait d'ailleurs déjà été rapporté dans l'étude de validation initiale du HOOS (31). D'autres études évaluant d'autres populations de coxarthroses sont donc nécessaires avant de pouvoir généraliser les résultats. Deuxièmement, l'intervalle réduit entre la première et la deuxième évaluation pour la sensibilité au changement (1 mois pour la viscosupplémentation, 3 mois pour la chirurgie) pourrait être jugé insuffisant. Cependant, les résultats en termes de taille de l'effet et de RMS ont été malgré tout satisfaisants.

Durant la dernière décennie, la traduction et l'adaptation transculturelle des instruments de mesure développés dans les pays anglophones est devenue une nécessité afin de pouvoir les utiliser dans les essais cliniques internationaux. Dans cette étude, la version française du HOOS a été obtenue puis validée selon les recommandations en vigueur dans le domaine de l'adaptation transculturelle des critères d'évaluation rapportés par le patient (18, 20, 22). Les propriétés psychométriques de cette version française sont bonnes et le HOOS paraît utile pour l'évaluation des patients atteints de coxarthrose, quel que soit le degré de sévérité de la coxarthrose. Bien que la cohérence interne du domaine *symptômes* nécessite d'être réévaluée, cette étude permet d'envisager l'utilisation de la version française du HOOS dans de futurs essais cliniques.

4. PROPRIETES PSYCHOMETRIQUES DES VERSIONS FRANÇAISES DU KOOS ET DU HOOS REDUITS (KOOS-PS ET HOOS-PS)

4.1 Résumé

Objectif. Evaluer les propriétés psychométriques de la version française du KOOS-PS (*Physical function Short form*) et HOOS-PS, en termes de faisabilité, de reproductibilité, de validité de construit et de sensibilité au changement.

Méthodes. Les patients qui consultaient pour une gonarthrose ou une coxarthrose dans le service de rhumatologie de Dijon ont été inclus. Lors de la première évaluation, les patients devaient compléter le questionnaire KOOS ou HOOS et le questionnaire AMIQUAL (Arthrose des Membres Inférieurs et QUALité de vie). Les patients devaient remplir un second questionnaire HOOS 15 jours plus tard pour étudier la reproductibilité par le coefficient de corrélation intra-classe (ICC). La faisabilité était évaluée par le pourcentage des réponses manquantes et la recherche d'effets plafond ou plancher. La validité de construit convergente et divergente était déterminée par la recherche de corrélations entre les résultats du KOOS-PS et du HOOS-PS et ceux de l'AMIQUAL, en utilisant le coefficient de corrélation de Spearman. La sensibilité au changement a été obtenue par le calcul de la réponse moyenne standardisée (RMS) et de la taille de l'effet, un mois après viscosupplémentation de hanche ou de genou.

Résultats. 87 patients avec gonarthrose et 50 patients avec coxarthrose ont été inclus. Le KOOS-PS et le HOOS-PS ont été obtenus chez tous les patients, en l'absence de réponse manquante. Aucun effet plancher ou plafond n'a été mis en évidence. Les ICCs du KOOS-PS et du HOOS-PS étaient respectivement égaux à 0.861 (0.763-0.921) et 0.859 (0.725-0.929). Une corrélation modérée à forte a été observée entre le KOOS-PS et le HOOS-PS et les

domaines *activités physiques*, *douleur* et *santé mentale* de l'AMIQUAL. Comme cela était attendu, une faible corrélation a été observée avec les autres domaines de l'AMIQUAL, sauf pour une corrélation modérée avec le domaine *activités sociales*. La sensibilité au changement était satisfaisante avec une RMS et une taille de l'effet respectivement de 0.80 et 0.51 pour le KOOS-PS et de 1.10 and 0.62 pour le HOOS-PS.

Conclusion. Les versions françaises du KOOS-PS et du HOOS-PS sont des questionnaires reproductibles, valides et sensibles au changement pour évaluer l'impotence fonctionnelle chez les patients avec une gonarthrose ou une coxarthrose.

4.2 Introduction

L'arthrose est la plus fréquente des maladies articulaires qui se caractérise par la destruction progressive du cartilage, et touche aussi bien les articulations portantes, comme le genou ou la hanche que les articulations des mains. La douleur et l'impotence fonctionnelle qui résultent de l'atteinte du genou ou de la hanche entraîne un retentissement important sur la qualité de vie des patients (225, 226). Comme la prévalence de la gonarthrose et la coxarthrose augmentent avec l'obésité, le vieillissement de la population et la diminution de l'activité physique, ces pathologies représenteront à l'avenir un problème majeur de santé publique. Il est donc important d'évaluer les traitements susceptibles de diminuer les répercussions de ces maladies et/ou susceptibles de ralentir ou de prévenir leur progression.

Divers instruments de mesure sont disponibles pour évaluer les capacités fonctionnelles chez les patients avec une gonarthrose ou une coxarthrose (23, 24, 225, 227). En particulier, le *Western Ontario and McMaster Universities Index* (WOMAC) est un auto-questionnaire largement utilisé et validé, qui permet d'évaluer la douleur la raideur et l'impotence fonctionnelle liées à l'arthrose (27). Comme il s'est avéré que l'échelle fonction du WOMAC

n'incluait pas des questions suffisamment difficiles, le KOOS (*Knee injury and Osteoarthritis Outcome Score*) et le HOOS (*Hip disability And Osteoarthritis Outcome Score*) ont été développés comme des extensions du WOMAC (28-31, 218, 231). Récemment, un groupe de travail mis en place sous les auspices de l'OARSI (*Osteoarthritis Research Society International*) et de l'OMERACT (*Outcome Measures in Rheumatology Clinical Trials*) a défini les différentes caractéristiques que devrait avoir un instrument de mesure pour évaluer de manière optimale la fonction chez des individus atteints de coxarthrose ou de gonarthrose, depuis le début de la maladie jusqu'à un stade avancé (220, 221). Parmi les principales considérations, ces questionnaires devraient évaluer l'ensemble du spectre de l'impotence fonctionnelle, sous une forme courte et avec des propriétés psychométriques suffisantes. En utilisant un modèle de Rasch et des données issues d'échantillons de patients représentant différents degrés de sévérité de la maladie, ce groupe d'experts a mis au point des instruments de mesure courts, capables de suivre la progression de l'impotence fonctionnelle dans la gonarthrose et la coxarthrose, appelés respectivement KOOS-PS et HOOS-PS (220, 221). Ces échelles de mesure réduites, dérivées du KOOS et du HOOS contiennent respectivement 7 questions pour le KOOS-PS et 5 pour le HOOS-PS. Ces outils de mesure, en couvrant un large niveau de difficulté et grâce à leur bonne faisabilité peuvent être intégrés pour évaluer la fonction dans des score composites de sévérité pour la gonarthrose et la coxarthrose.

Du fait du nombre croissant d'études multicentriques internationales et de la nécessité d'obtenir des résultats significatifs pour les études épidémiologiques et/ou thérapeutiques, il est indispensable de réaliser l'adaptation et la validation transculturelle des instruments de mesure de la santé. Il est par ailleurs indispensable d'évaluer les propriétés psychométriques de tout instrument de mesure, comme défini par le filtre OMERACT (49) en étudiant sa validité (refléter ce qu'il est censé mesurer), sa reproductibilité et sa sensibilité au changement. Le dernier critère du filtre OMERACT correspond à la faisabilité de l'instrument

de mesure, qui prend en compte le temps nécessaire, la disponibilité, le coût et n'est pas analysable de manière statistique.

L'objectif de cette étude était d'évaluer les propriétés psychométriques de la version française du KOOS-PS et du HOOS-PS en termes de faisabilité, reproductibilité, validité de construit et sensibilité au changement.

4.3 Méthodes

Design de l'étude : étude prospective mono-centrique

Patients

Les patients qui consultaient de manière consécutive pour une gonarthrose ou une coxarthrose symptomatique dans le service de rhumatologie de l'hôpital universitaire de Dijon ont été inclus. Les critères d'inclusion pour les patients étaient les suivants : âge supérieur à 40 ans, gonarthrose ou coxarthrose primitive répondant aux critères de l'*American College of Rheumatology* (233, 236), capacité à comprendre les ordres simples et à compléter seuls les auto-questionnaires. Pour les patients évalués pour la sensibilité au changement, un critère d'inclusion supplémentaire était l'indication retenue par le rhumatologue traitant de viscosupplémentation par acide hyaluronique. Les critères d'exclusion étaient définis par la présence d'une autre maladie rhumatologique invalidante comme des lombalgies chroniques ou une autre localisation arthrosique au membre inférieur, une arthrose en poussée inflammatoire à l'examen clinique, ou la réalisation d'une infiltration intra-articulaire de corticoïde au cours des trois derniers mois. Chez les patients évalués pour l'étude de la reproductibilité, un critère d'exclusion supplémentaire était un changement dans le traitement de l'arthrose après la première évaluation. Chez les patients évalués pour l'étude de la

sensibilité au changement, un critère d'exclusion supplémentaire était un changement dans le traitement de l'arthrose dans le mois à venir, hormis la viscosupplémentation.

Questionnaires

Lors de la première évaluation, les patients devaient répondre aux versions françaises du KOOS et du HOOS dont le processus d'adaptation transculturelle a été décrit auparavant (80, 240). En résumé, trois personnes (deux rhumatologues et un professeur d'anglais d'origine française) ont traduit la version anglaise de ces questionnaires. Après une réunion entre traducteurs, une seule version a été obtenue. Cette version a été rétro-traduite par un traducteur bilingue, anglais d'origine, en aveugle de la version originale. Une réunion multidisciplinaire a permis d'aboutir par consensus à la version finale, après avoir vérifié si son contenu était compréhensible et si les équivalences transculturelles avaient été respectées. Lors de la dernière étape, la version finale a été pré-testée chez 15 patients avec une gonarthrose et 15 patients avec une coxarthrose.

Le KOOS-PS comprend les 7 questions suivantes: sortir du lit, se pencher en avant, mettre ses chaussettes ou ses collants, se relever d'une position assise, tourner ou pivoter sur sa jambe, rester à genoux, s'accroupir et le HOOS-PS en contient 5 : descendre les escaliers, entrer ou sortir d'une baignoire, rester assis, courir, tourner ou pivoter sur votre jambe, ce qui permet de calculer leurs scores respectifs. Les scores initiaux (de 0 à 28 pour les KOOS-PS et de 0 à 20 pour le HOOS-PS) ont été normalisés entre 0 et 100, 0 indiquant le meilleur score (220, 221). Les patients évalués pour l'étude de la validité devaient aussi compléter le questionnaire AMIQUAL (Arthrose des Membres Inférieurs et QUALité de vie) lors de la première évaluation. L'AMIQUAL est un questionnaire spécifique de la qualité de vie dans l'arthrose des membres inférieurs qui a été récemment validé (179). L'AMIQUAL contient 43 questions définissant 5 domaines (*douleur, activités physiques, santé mentale, soutien social,*

activités sociales) et 3 questions indépendantes (activité sexuelle, activités professionnelles). Les scores vont de 0 (retentissement maximal) à 100 (retentissement nul). Les patients évalués pour la reproductibilité devaient remplir un second questionnaire KOOS ou HOOS 15 jours plus tard, qu'ils devaient renvoyer par courrier, en utilisant une enveloppe préaffranchie. Cet intervalle de temps a été jugé suffisamment long pour s'assurer de l'absence de mémorisation des réponses et suffisamment court pour éviter un changement significatif d'état lié à la gonarthrose ou à la coxarthrose. Les patients évalués pour la sensibilité au changement étaient traités par viscosupplémentation par acide hyaluronique. Les patients avec une coxarthrose bénéficiaient d'une injection d'acide hyaluronique sous guidage échographique. L'indication de viscosupplémentation était retenue par le rhumatologue traitant mais l'injection était systématiquement réalisée par le même clinicien (PO). La nature et le poids moléculaire de l'acide hyaluronique variaient puisque les patients venaient avec la prescription réalisée par leur rhumatologue traitant. Les patients avec une gonarthrose bénéficiaient de trois viscosupplémentations non échoguidées, réalisées à une semaine d'intervalle par leur rhumatologue traitant, selon le protocole habituel. La nature et le poids moléculaire de l'acide hyaluronique variait donc aussi. Les patients devaient remplir un deuxième questionnaire KOOS ou HOOS un mois après la dernière viscosupplémentation et devaient le renvoyer par courrier dans une enveloppe préaffranchie.

Pour le KOOS-PS et le HOOS-PS quand au moins une réponse était manquante, le score n'était pas calculé. Pour l'AMIQUAL, quand plus de la moitié des réponses d'un domaine étaient manquantes, le score n'était pas calculé. Quand moins de la moitié des réponses d'un domaine étaient manquantes, elles étaient remplacées par la moyenne des réponses obtenues pour ce domaine chez chaque individu.

Analyse statistique

Faisabilité

La faisabilité a été évaluée par le pourcentage de réponses manquantes et par la recherche d'effet plancher ou plafond, considérés comme présents si plus de 15% des patients obtenaient le score maximum ou minimum pour le KOOS-PS ou le HOOS-PS.

Reproductibilité

La reproductibilité du KOOS-PS et du HOOS-PS a été évaluée en utilisant les réponses aux questionnaires remplis à 15 jours d'intervalle, en calculant le coefficient de corrélation intra-classe (ICC), avec son intervalle de confiance à 95%. Un ICC supérieur à 0.8 est habituellement considéré comme le témoin d'une excellente reproductibilité. Par ailleurs, la méthode graphique de Bland et Altman, dans laquelle la différence entre la première et la deuxième évaluation est représentée en fonction de la moyenne des deux évaluations a été obtenue pour chaque questionnaire. Ce type de représentation permet de décrire le pourcentage de patients et leur distribution à l'intérieur des bornes de l'intervalle de confiance à 95% pour la différence entre les mesures répétées.

Validité de construit

La validité de construit (convergente et divergente) a été déterminée en comparant les résultats du KOOS-PS et du HOOS-PS à ceux de l'AMIQUAL. Le coefficient de corrélation de Spearman a été utilisé pour évaluer les corrélations entre ces questionnaires. Des coefficients de corrélation >0.5 , entre 0.35-0.5, et <0.35 ont été considérés respectivement comme des corrélations forte, modérée et faible (82). Des hypothèses ont été formulées a priori pour la validité convergente (corrélation modérée à forte) et divergente (faible corrélation) selon que les instruments mesuraient des concepts proches ou éloignés et selon

les données issues de la littérature. Il a ainsi été supposé que le KOOS-PS et le HOOS-PS seraient modérément ou fortement corrélés aux domaines *douleur* et *activités physiques* de l'AMIQUAL, mais aussi au domaine *santé mentale* car il a été démontré que ce domaine particulier était fortement corrélé au domaine fonction du WOMAC (80). A l'inverse, il a été supposé que le KOOS-PS et le HOOS-PS seraient faiblement corrélés avec les autres domaines de l'AMIQUAL (*soutien social* et *activités sociales*).

Sensibilité au changement

La sensibilité au changement a été évaluée en comparant les résultats avant et 1 mois après la viscosupplémentation par acide hyaluronique. La réponse moyenne standardisée (RMS), qui correspond à la différence moyenne entre les scores avant et après traitement divisée par l'écart type de cette différence, et la taille de l'effet, qui correspond à la différence moyenne entre les scores avant et après traitement divisé par l'écart-type du score avant traitement, ont ainsi été calculées.

Le logiciel SPSS (*Statistical Package for the Social Sciences* (SPSS) version 14.0) a été utilisé pour l'analyse statistique. Le degré de significativité statistique était défini pour un $p < 0.05$.

4.4 Résultats

49 patients avec une gonarthrose (âge moyen = 72 ± 9 ans, 71% femmes) ont été inclus dans l'étude de la reproductibilité dont 36 participèrent aussi à l'étude de la validité de construit. Parmi ces 49 patients, 46 ont renvoyé leur questionnaire 15 jours après. La sensibilité au changement a été évaluée chez 38 autres patients, qui ont tous retourné leur second questionnaire un mois après la viscosupplémentation de genou.

30 patients avec une coxarthrose (âge moyen 65 ± 10 ans, 74% femmes) ont été inclus dans l'étude de la reproductibilité et de la validité de construit. Tous ont renvoyé leur second questionnaire 15 jours après. La sensibilité aux changements a été évaluée chez 20 autres patients qui ont tous renvoyé leur second questionnaire un mois après la viscosupplémentation de hanche.

Comme aucune des réponses du KOOS-PS et du HOOS-PS n'était manquante parmi l'ensemble des questionnaires recueillis, les scores ont pu être calculés chez tous les patients. Mis à part les trois questions indépendantes concernant les répercussions professionnelles (la plupart des patients étant à la retraite) et sexuelles qui n'entrent pas en compte pour le calcul du score, le reste du questionnaire AMIQUAL a été rempli correctement, ce qui a permis de calculer les scores des différents domaines chez tous les patients. Sachant qu'aucun patient n'a atteint le score minimum ou maximum pour le KOOS-PS et le HOOS-PS, aucun effet plancher ou plafond n'a été observé, quel que soit le moment du remplissage. La reproductibilité du KOOS-PS et du HOOS-PS était excellente (tableau 1).

	1^{ère} Evaluation moyenne (DS)	2^{ème} Evaluation moyenne (DS)	ICC (IC 95%)
<i>KOOS-PS</i>	48.3 (12.8)	49.1 (12.2)	0.861 (0.763-0.921)
<i>HOOS-PS</i>	51.1 (7.7)	52.1 (18)	0.859 (0.725-0.929)

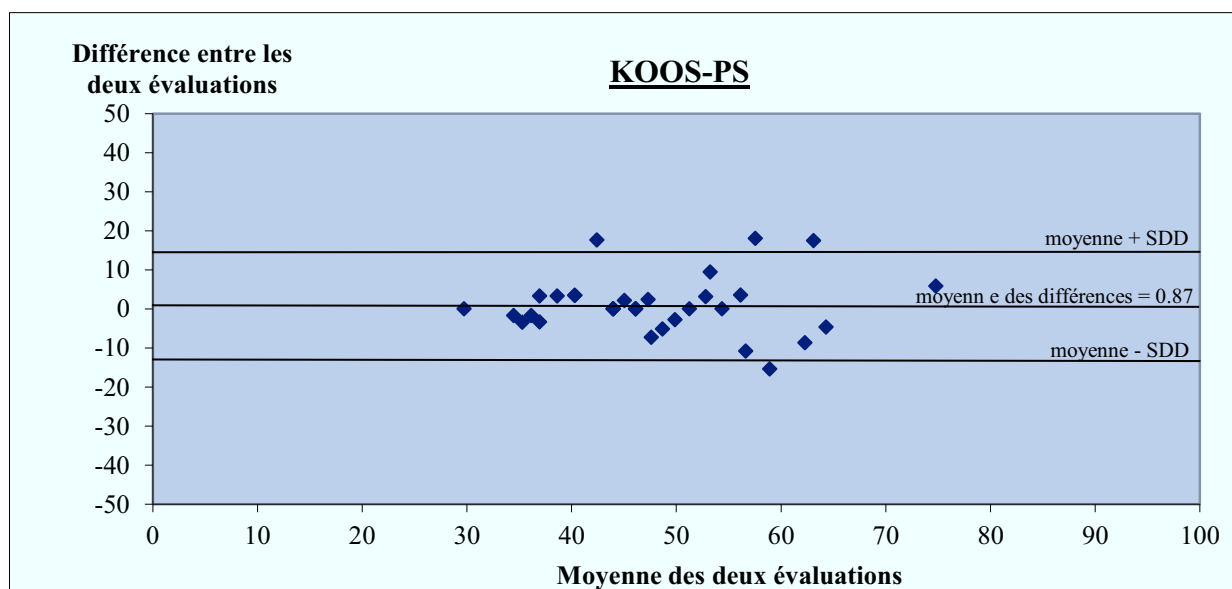
Tableau 1. Scores moyens du KOOS-PS et du HOOS-PS réalisés à 15 jours d'intervalle et reproductibilité de mesure

ICC : coefficient de corrélation intra-classe DS : déviation standard
IC 95% : intervalle de confiance à 95%

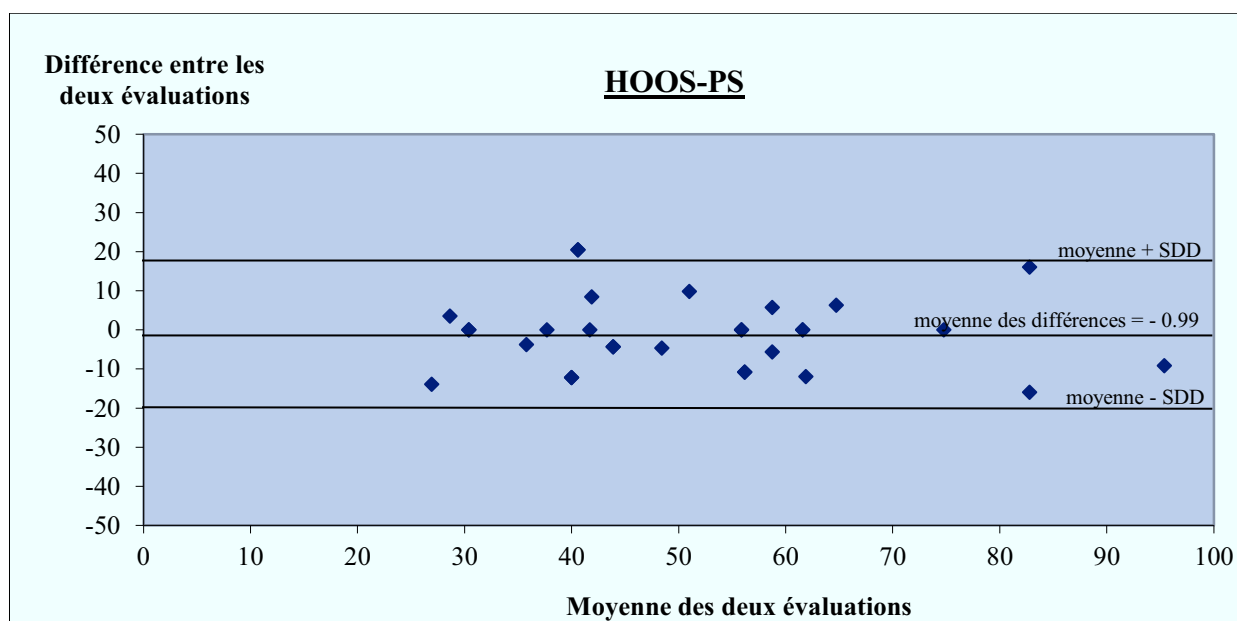
La méthode graphique de Bland et Altman est présentée dans la figure 1. Les différences entre les mesures répétées étaient incluses dans les bornes de l'intervalle de confiance dans la majorité des cas et n'étaient pas liées à la moyenne entre les deux évaluations.

Figure. 1. Reproductibilité des versions françaises du KOOS-PS (**Figure 1a**) et du HOOS-PS (**Figure 1b**) par la méthode graphique de Bland et Altman. Les deux évaluations ont été réalisées à 15 jours d'intervalle

▪ **Figure 1a.**



▪ **Figure 1b.**



Les résultats de la validité de construit (convergente et divergente) sont présentés dans le tableau 2. Ils montrent que 8 des 10 hypothèses formulées a priori ont été confirmées.

AMIQUAL	KOOS-PS	HOOS-PS
<i>Activités physiques</i>	-0.44 (p = 0.009)	-0.665 (p < 0.001)
<i>Douleur</i>	-0.366 (p = 0.033)	-0.385 (p = 0.036)
<i>Santé mentale</i>	-0.328 (p = 0.06)	-0.473 (p < 0.001)
<i>Soutien social</i>	-0.03 (p = 0.87)	0.022 (p = 0.91)
<i>Activités sociales</i>	-0.34 (p = 0.05)	-0.125 (p = 0.5)

Tableau 2. Validité de construit : corrélations entre le KOOS-PS et HOOS-PS et le questionnaire AMIQUAL (Test de Spearman)

Comme attendu, une corrélation modérée à forte a été observée entre le KOOS-PS, le HOOS-PS et les domaines *activités physiques*, *douleur* et *santé mentale* de l'AMIQUAL. Une faible corrélation a été observée entre le KOOS-PS, le HOOS-PS et les autres domaines de l'AMIQUAL, à l'exception d'une corrélation modérée entre le KOOS-PS et le domaine *activités sociales*.

La sensibilité au changement montrait des effets importants (tableau 3, page suivante) avec une RMS et une taille de l'effet respectivement de 0.80 et 0.51 pour le KOOS-PS et de 1.1 et 0.62 pour le HOOS-PS. Après viscosupplémentation, les scores du KOOS-PS et du HOOS-PS s'amélioraient chez une grande majorité de patients (respectivement, 33/38 et 17/20).

	Avant viscosupplémentation <i>moyenne (DS)</i>	Après visco-supplémentation <i>moyenne (DS)</i>	RMS	Taille de l'effet
KOOS-PS	48.2 (15.6)	40.3 (13.9)	0.80	0.51
HOOS-PS	51.3 (16.2)	41.3 (17.4)	1.10	0.62

Tableau 3. Sensibilité au changement des versions françaises du KOOS-PS et du HOOS-PS. Les patients ont été évalués avant et un mois après viscosupplémentation.
RMS : Réponse moyenne standardisée
DS : déviation standard

4.5 Discussion

Dans cette étude, les propriétés psychométriques du KOOS-PS et du HOOS-PS ont été évaluées et apparaissent satisfaisantes. Toutefois, ces résultats devront d'être confirmés par d'autres études évaluant d'autres populations de patients atteints d'arthrose, avant de pouvoir être généralisés. Par ailleurs, les données utilisées dans cette étude ont été extraites depuis les versions complètes des questionnaires KOOS et HOOS. Il est nécessaire de confirmer ces résultats chez des patients remplissant spécifiquement les questionnaires KOOS-PS et HOOS-PS.

L'étude des corrélations avec l'AMIQAL montrent une validité convergente et divergente satisfaisante pour le HOOS-PS et le KOOS-PS. Comme attendu, de fortes corrélations ont été mises en évidence entre les domaines de l'AMIQAL et du KOOS-PS et du HOOS-PS qui mesuraient les mêmes concepts (activités physiques). Les bonnes corrélations avec le domaine *douleur* de l'AMIQAL sont concordants avec ceux de précédentes études qui avaient démontré le lien entre la fonction et la douleur chez les patients

arthrosiques (231, 235). La corrélation modérée entre le KOOS-PS et le HOO-PS et le domaine *santé mentale* de l'AMIQUAL était attendue car des corrélations avaient été retrouvées précédemment entre ce domaine particulier de l'AMIQUAL et les domaines douleur et fonction du WOMAC (80).

Le domaine *santé mentale* de l'AMIQUAL était statistiquement corrélé au HOOS-PS mais pas au KOOS-PS. Toutefois, cette dernière corrélation atteignait presque la significativité ($p = 0.06$), ce qui suggère que les deux versions réduites sont corrélées avec ce domaine et que la différence observée est due aux fluctuations d'échantillonnage. Une corrélation modérée a été retrouvée entre le KOOS-PS et le domaine *activités sociales* de l'AMIQUAL. Ceci pourrait être dû à l'inclusion dans ce domaine de l'AMIQUAL de deux questions (« je sors de chez moi autant que je le souhaite », « je reçois chez moi autant que je le souhaite ») qui peuvent avoir des rapports avec les capacités fonctionnelles du patient. Cependant, aucune corrélation n'a été observée entre ce domaine *activités sociales* de l'AMIQUAL et le HOOS-PS. D'autres études sont nécessaires pour savoir si ces résultats discordants sont liés à la fluctuation d'échantillonnage ou s'ils sont réels.

La reproductibilité des versions réduites du HOOS et du KOOS était excellente, en particulier en considérant l'intervalle de temps entre les deux mesures. La méthode graphique de Bland et Altman a permis de vérifier que les différences entre les mesures répétées n'étaient pas liées à la moyenne de ces mesures. La sensibilité au changement était bonne ou excellente. Elle pourrait même être considérée comme étonnamment élevée en regard du traitement utilisé, particulièrement pour la hanche car certaines études n'ont pas réussi à démontrer la supériorité de l'acide hyaluronique sur le placebo dans la coxarthrose (242, 247). L'amélioration pourrait être liée à un effet placebo, qui est connu comme étant particulièrement important dans l'arthrose, en particulier pour les traitements injectables (245). Toutefois, l'objectif de cette étude n'était pas d'évaluer l'efficacité de la

viscosupplémentation par acide hyaluronique dans le gonarthrose ou la coxarthrose mais d'évaluer la sensibilité au changement du questionnaire. De plus, alors que le KOOS-PS et le HOOS-PS avaient démontré auparavant leur capacité à détecter l'amélioration fonctionnelle après arthroplastie (224), il est intéressant de constater que ces questionnaires peuvent aussi détecter des changements plus minimes pour des traitements à l'efficacité moins importante. Il convient de souligner que l'intervalle de 1 mois pour la seconde évaluation était suffisant car l'amélioration après viscosupplémentation apparaît 15 jours après l'injection (248).

Mis à part la reproductibilité, la validité et la sensibilité au changement, une autre propriété psychométrique cruciale d'un questionnaire pour les essais cliniques est sa faisabilité. Aucun effet plancher ou plafond n'a été observé avec les versions réduites des questionnaires. Une limite potentielle du KOOS-PS et du HOOS-PS est que, contrairement à leurs versions complètes qui autorisent jusqu'à deux données manquantes par domaine (29), ils nécessitent une réponse à chacune des questions. L'absence de données manquantes dans cette étude est donc rassurante à ce sujet.

Le KOOS-PS et le HOOS-PS ont été développés par un groupe de travail sous l'égide de l'OARSI et de l'OMERACT afin d'établir au final un critère composite pour l'indication de pose de prothèse de hanche, qui serait utilisé comme critère de jugement dur dans les essais thérapeutiques (222). Pour parvenir à cet objectif, il est indispensable d'utiliser des instruments qui soient capables d'évaluer la progression de l'impotence fonctionnelle depuis le début de la maladie jusqu'à un stade avancé. Les avantages théoriques du KOOS-PS et du HOOS-PS sont aussi d'éviter la redondance entre certaines questions. Le modèle de Rasch utilisé permet en outre de créer une échelle de mesure évaluant différents niveaux de difficulté pour les capacités fonctionnelles, sans être dépendants les uns des autres. Comme ces deux questionnaires réduits ont été développés sur la base du KOOS et du HOOS qui ont de nombreuses questions en commun, ils pourraient ne pas être capables de distinguer la

gonarthrose de la coxarthrose. En effet, les 5 questions du HOOS-PS sont comprises dans le KOOS et 6 des 7 questions du KOOS-PS sont comprises dans le HOOS. Cependant, la finalité de ces questionnaires réduits n'est pas de discriminer les deux maladies. De plus, une seule question est finalement commune aux deux questionnaires réduits, ce qui signifie que les autres questions du KOOS-PS ne sont pas applicables au modèle de Rasch pour les patients avec une coxarthrose et que les autres questions du HOOS-PS ne sont pas applicables au modèle de Rasch des patients avec gonarthrose. Ceci suggère que ces versions réduites sont donc capables de distinguer les deux maladies.

Une précédente étude du HOOS-PS et KOOS-PS chez des patients avec indication de prothèse totale de hanche ou de genou avait déjà mis en évidence la validité de construit et la sensibilité au changement de ces questionnaires, en les comparant au WOMAC fonction (224). Toutefois, à notre connaissance, la reproductibilité, la validité de construit chez des patients avec une arthrose moins sévère et la sensibilité au changement de ces instruments après un traitement moins efficace que l'arthroplastie n'avaient jamais été encore évaluées. Dans cette étude, les versions françaises du KOOS-PS et du HOOS-PS ont démontré des résultats satisfaisants en termes de faisabilité, de validité de construit convergente et divergente, de reproductibilité et de sensibilité au changement. Par ailleurs, ce travail peut être considéré comme une étude de validation supplémentaire de ces nouveaux instruments de mesure pour la gonarthrose et la coxarthrose qui visent à évaluer les répercussions fonctionnelles de ces pathologies.

A l'issue de ce travail, il nous est apparu opportun de proposer la même évaluation méthodologique du KOOS, du HOOS, du KOOS-PS et du HOOS-PS que celle déjà présentée en deuxième partie de ce travail, en y incluant les données provenant des études de validation française présentées ci-dessus.

5. Synthèse de l'évaluation méthodologique des questionnaires KOOS et HOOS

5.1 KOOS

Knee injury and Osteoarthritis Outcome score (28, 29, 82, 216, 218, 231, 232, 240, 249-255)					
Etude de validation initiale		Roos et al (J Orthop Sports Phys Ther, 1998)			
Pathologie(s) cible(s)		Traumatismes du genou et gonarthrose			
Version transculturelle publiée		Anglais - Allemand - Portugais - Français - Hollandais - Chinois - Iranien			
Questions	Nombre	42			
	Score	0 à 4			
Domaines	Nombre	5			
	Dénomination	Douleur - Symptômes - Fonction activités quotidiennes - Fonction activités sports et loisirs - Qualité de vie			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Lysholm				
	WOMAC				
	EVA douleur				
	EQ-5D				
	SF-12				
	OAKHQOL				
Sensibilité au changement	Prothèse totale				
	Viscosupplém.				
	Chir ménisques				
	Chir cartilage				
	Chir lig croisés				
Erreur de Mesure					

5.2 KOOS-PS

Knee injury and Osteoarthritis Outcome Score - Physical function Short form (79, 220, 224, 256)					
Etude de validation initiale		Perruccio A.V et al (Osteoarthritis and cartilage, 2008)			
Pathologie(s) cible(s)		Traumatismes du genou et gonarthrose			
Validation transculturelle publiée		Anglais - Français			
Questions	Nombre	7			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Chronic pain grade				
	HADS				
	Fatigue POMS				
	OAKHQOL				
Sensibilité au changement	Prothèse totale				
	Viscosupplém.				
Erreur de mesure					

5.3 HOOS

Hip disability and Osteoarthritis Outcome Score (30, 31, 80, 156, 219)					
Etude de validation initiale		Klassbo M et al (Scan J Rheum, 2003)			
Pathologie(s) cible(s)		Pathologies de hanche dont coxarthrose			
Validation transculturelle publiée		Anglais - Français - Hollandais - Allemand - Lituanien - Danois - Suédois			
Questions	Nombre	40			
	Score	0 à 4			
Domaines	Nombre	5			
	Dénomination	Douleur - Symptômes - Fonction activités quotidiennes - Fonction activités sports et loisirs - Qualité de vie			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement maximal 100 = retentissement nul			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	SF-36				
	Oxford score				
	EVA douleur				
	Lequesne				
Sensibilité au changement	Prothèse totale				
	Viscosupplém.				
Erreur de mesure					

5.4 HOOS-PS

Hip disability and Osteoarthritis Outcome Score -Physical function Short form (79, 221, 224)					
Etude de validation initiale		Davis A et al (Osteoarthritis and cartilage, 2008)			
Pathologie(s) cible(s)		Pathologies de hanche dont coxarthrose			
Validation transculturelle publiée		Anglais - Français			
Questions	Nombre	5			
	Score	0 à 4			
Domaines	Nombre	1			
	Dénomination	Fonction			
Score total	Valeur	0 à 100			
	Interprétation	0 = retentissement nul 100 = retentissement maximal			
Propriétés psychométriques		Non évaluée	Évaluée		
			-	±	+
Faisabilité	% réponses				
	Plafond/plancher				
	Tps remplissage				
	Tps cotation				
Reproductibilité					
Cohérence interne					
Validité de construit	Chronic pain grade				
	HADS				
	Fatigue POMS				
	OAKHQOL				
Sensibilité au changement	Prothèse totale				
	Viscosupplém.				
Erreur de mesure					

Conclusion

La validation transculturelle des auto-questionnaires, problématique centrale de ce travail de thèse, représente désormais une étape incontournable pour tous les PROs susceptibles d'être utilisés dans les essais thérapeutiques (5, 6). Grâce aux efforts conjugués des autorités de régulation et des sociétés savantes, de larges progrès ont été effectués ces dernières années (42), comme l'illustrent nos travaux, pour relever le challenge de la validation transculturelle. Cependant, la validation transculturelle ne représente qu'une étape dans le long cheminement nécessaire pour développer et valider un nouveau questionnaire (5). La complexité de ce processus de validation transculturelle nécessite l'intervention d'équipes multidisciplinaires et spécialisées. Les résultats des travaux de ces équipes doivent être publiés dans des revues référencées pour faciliter leur diffusion (21) et préciser la méthodologie employée. Certaines versions traduites d'auto-questionnaires, comme par exemple la version française du WOMAC, n'ont jamais bénéficié d'une véritable adaptation transculturelle. Ceci est particulièrement préjudiciable étant donnée la notoriété rapidement acquise par ce questionnaire, sa large utilisation dans les essais cliniques internationaux et son rôle comme critère de référence pour la validité convergente des versions françaises de PROs dans le domaine de l'arthrose.

Ce travail de thèse a permis la validation transculturelle des versions françaises du KOOS, DU HOOS et de leurs formes réduites. Les résultats de ce travail permettent en outre d'apporter de nouvelles données sur la validité de ces questionnaires. Ainsi, pour la première fois, la sensibilité au changement pour des traitements intra-articulaires injectables (acide hyaluronique) a été évaluée dans une cohorte de patients arthrosiques. Les résultats sont bons et, même avec les limites de ces études, sont particulièrement intéressants pour évaluer des patients qui ont souvent des atteintes mono-articulaires. En effet, très peu de données sont disponibles sur la validation de PROs pour des traitements intra-articulaires (209) alors que de nombreux traitements sont en développement à ce niveau dans l'arthrose (257) ou les

rhumatismes inflammatoires (258). Ces travaux, tout comme ceux accomplis par notre équipe lors de la validation française du RAOS avec les infiltrations de corticoïdes pour les monoarthropathies inflammatoires dans la polyarthrite rhumatoïde (185), répondent ainsi parfaitement à la demande récente de l'OMERACT pour une meilleure évaluation de ces traitements intra-articulaires (209).

Parallèlement, nos travaux ont aussi permis d'évaluer la validité de construit du KOOS, du HOOS, du KOOS-PS et du HOOS-PS avec des questionnaires spécifiques de l'arthrose des membres inférieurs. Ainsi leurs corrélations avec l'indice de Lequesne, qui est l'une des références pour l'évaluation des répercussions fonctionnelles de la gonarthrose et de la coxarthrose, et avec l'AMIQUAL, qui est à ce jour le seul questionnaire spécifique de la qualité de vie dans l'arthrose des membres inférieurs, se sont avérées satisfaisantes. D'autres études de validation restent nécessaires pour confirmer ces résultats dans des populations différentes, pour évaluer la sensibilité au changement avec d'autres traitements et ainsi mieux préciser l'intérêt du KOOS et du HOOS par rapport au gold standard, à savoir le WOMAC.

Grâce aux travaux initiaux remarquables de l'équipe d'Ewa Roos, le KOOS et le HOOS sont des instruments de mesure amenés à devenir des PROs de référence dans les essais cliniques. Dans le domaine de l'arthrose, en plus de leur excellente validité de contenu favorisée par l'intégration du WOMAC, ils capturent des concepts supplémentaires par rapport à ce dernier (retentissement fonctionnel dans les activités sportives en particulier). Ils ont en outre démontré une meilleure sensibilité au changement et une capacité discriminante supérieure. Ces questionnaires ne peuvent en revanche être considérés comme des outils de référence pour l'évaluation de la qualité de vie dans l'arthrose, comme le sont l'AMIQUAL ou l'AIMS, car ils ne l'abordent que très partiellement à travers quatre questions.

La difficulté du passage à la pratique quotidienne des PROs évaluant la fonction et/ou la qualité de vie a aussi été soulevée dans cette thèse (2). En effet, les 42 questions du KOOS et les 40 du HOOS compromettent sa faisabilité en pratique clinique, d'autant plus que le calcul des scores n'est pas simple. Le développement et la validation transculturelle de formes réduites pour le KOOS et pour le HOOS permettent d'envisager une utilisation en pratique quotidienne dans l'avenir mais ceci reste encore à évaluer. Ces formes réduites, le KOOS-PS et le HOOS-PS, ont dès à présent un intérêt évident en recherche épidémiologique. Elles peuvent en effet s'intégrer dans des scores composites, en plein développement en Rhumatologie (259, 260). A ce titre, l'OMERACT et l'OARSI ont retenu ces formes réduites du KOOS et du HOOS pour évaluer la fonction dans une importante étude internationale pour la mise au point d'un indice composite visant à définir le recours à la prothèse totale dans l'arthrose des membres inférieurs. Ceci est un signe fort de la reconnaissance du travail accompli ces dernières années pour l'ensemble des intervenants qui ont participé à travers le monde au processus de développement et de validation de ces questionnaires.

Bibliographie

1. DOUGADOS, M. (2004) Monitoring osteoarthritis progression and therapy, *Osteoarthritis Cartilage*, 12 Suppl A, S55-60.
2. GUILLEMIN, F. (2000) Functional disability and quality-of-life assessment in clinical practice, *Rheumatology (Oxford)*, 39 Suppl 1, 17-23.
3. DOWARD, L. C. & MCKENNA, S. P. (2004) Defining patient-reported outcomes, *Value Health*, 7 Suppl 1, S4-8.
4. PINCUS, T., YAZICI, Y. & BERGMAN, M. J. (2009) Patient questionnaires in rheumatoid arthritis: advantages and limitations as a quantitative, standardized scientific medical history, *Rheum Dis Clin North Am*, 35, 735-43, vii.
5. (2009) U.S. Food and Drug Administration. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.
6. (2005) Committee for Medicinal Products for Human Use (CHMP). Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. London: European Medicines Agency (EMA).
7. PINCUS, T., YAZICI, Y. & SOKKA, T. (2007) Quantitative measures of rheumatic diseases for clinical research versus standard clinical care: differences, advantages and limitations, *Best Pract Res Clin Rheumatol*, 21, 601-28.
8. KIRWAN, J. R., NEWMAN, S., TUGWELL, P. S. et al. (2009) Progress on incorporating the patient perspective in outcome assessment in rheumatology and the emergence of life impact measures at OMERACT 9, *J Rheumatol*, 36, 2071-6.
9. GUYATT, G. H., FEENY, D. H. & PATRICK, D. L. (1993) Measuring health-related quality of life, *Ann Intern Med*, 118, 622-9.
10. TERWEE, C. B., BOT, S. D., DE BOER, M. R. et al. (2007) Quality criteria were proposed for measurement properties of health status questionnaires, *J Clin Epidemiol*, 60, 34-42.
11. PATRICK, D. L., BURKE, L. B., POWERS, J. H. et al. (2007) Patient-reported outcomes to support medical product labeling claims: FDA perspective, *Value Health*, 10 Suppl 2, S125-37.
12. KIRWAN, J. R., NEWMAN, S., TUGWELL, P. S. & WELLS, G. A. (2009) Patient perspective on outcomes in rheumatology -- a position paper for OMERACT 9, *J Rheumatol*, 36, 2067-70.
13. BELLAMY, N., KIRWAN, J., BOERS, M. et al. (1997) Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III, *J Rheumatol*, 24, 799-802.
14. BOMBARDIER, C. & TUGWELL, P. (1987) Methodological considerations in functional assessment, *J Rheumatol Suppl*, 14 Suppl 15, 6-10.
15. FITZPATRICK, R., DAVEY, C., BUXTON, M. J. & JONES, D. R. (1998) Evaluating patient-based outcome measures for use in clinical trials, *Health Technol Assess*, 2, i-iv, 1-74.
16. GUYATT, G. H., KIRSHNER, B. & JAESCHKE, R. (1992) Measuring health status: what are the necessary measurement properties?, *J Clin Epidemiol*, 45, 1341-5.
17. HUNT, S. M., ALONSO, J., BUCQUET, D. et al. (1991) Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment, *Health Policy*, 19, 33-44.
18. GUILLEMIN, F. (1995) Cross-cultural adaptation and validation of health status measures, *Scand J Rheumatol*, 24, 61-3.
19. HERDMAN, M., FOX-RUSHBY, J. & BADIA, X. (1997) 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires, *Qual Life Res*, 6, 237-47.
20. GUILLEMIN, F., BOMBARDIER, C. & BEATON, D. (1993) Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines, *J Clin Epidemiol*, 46, 1417-32.
21. BEATON, D. E., BOMBARDIER, C., GUILLEMIN, F. & FERRAZ, M. B. (2000) Guidelines for the process of cross-cultural adaptation of self-report measures, *Spine (Phila Pa 1976)*, 25, 3186-91.
22. WILD, D., GROVE, A., MARTIN, M. et al. (2005) Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation, *Value Health*, 8, 94-104.

23. GARRATT, A. M., BREALEY, S. & GILLESPIE, W. J. (2004) Patient-assessed health instruments for the knee: a structured review, *Rheumatology (Oxford)*, 43, 1414-23.
24. MARX, R. G. (2003) Knee rating scales, *Arthroscopy*, 19, 1103-8.
25. THORBORG, K., ROOS, E. M., BARTELS, E. M., PETERSEN, J. & HOLMICH, P. (2009) Validity, Reliability And Responsiveness Of Patient-Reported Outcome Questionnaires When Assessing Hip And Groin Disability: A Systematic Review, *Br J Sports Med*.
26. TERWEE, C. B., MOKKINK, L. B., STEULTJENS, M. P. & DEKKER, J. (2006) Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties, *Rheumatology (Oxford)*, 45, 890-902.
27. BELLAMY, N., BUCHANAN, W. W., GOLDSMITH, C. H., CAMPBELL, J. & STITT, L. W. (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee, *J Rheumatol*, 15, 1833-40.
28. ROOS, E. M., ROOS, H. P., LOHMANDER, L. S., EKDAHL, C. & BEYNNON, B. D. (1998) Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure, *J Orthop Sports Phys Ther*, 28, 88-96.
29. ROOS, E. M., ROOS, H. P., EKDAHL, C. & LOHMANDER, L. S. (1998) Knee injury and Osteoarthritis Outcome Score (KOOS)--validation of a Swedish version, *Scand J Med Sci Sports*, 8, 439-48.
30. KLASSBO, M., LARSSON, E. & MANNEVIK, E. (2003) Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index, *Scand J Rheumatol*, 32, 46-51.
31. NILSDOTTER, A. K., LOHMANDER, L. S., KLASSBO, M. & ROOS, E. M. (2003) Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement, *BMC Musculoskelet Disord*, 4, 10.
32. DOUGADOS, M., HAWKER, G., LOHMANDER, S. et al. (2009) OARSI/OMERACT criteria of being considered a candidate for total joint replacement in knee/hip osteoarthritis as an endpoint in clinical trials evaluating potential disease modifying osteoarthritic drugs, *J Rheumatol*, 36, 2097-9.
33. VALLERAND, R. J. & HALLIWELL, W. R. (1983) [One method of trans-cultural validation of psychological questionnaires: implications for sports psychology], *Can J Appl Sport Sci*, 8, 9-18.
34. GJERSING, L., CAPLEHORN, J. R. & CLAUSEN, T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations, *BMC Med Res Methodol*, 10, 13.
35. DA MOTA FALCAO, D., CICONELLI, R. M. & FERRAZ, M. B. (2003) Translation and cultural adaptation of quality of life questionnaires: an evaluation of methodology, *J Rheumatol*, 30, 379-85.
36. BOWDEN, A. & FOX-RUSHBY, J. A. (2003) A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America, *Soc Sci Med*, 57, 1289-306.
37. BEACH, M. C., GARY, T. L., PRICE, E. G. et al. (2006) Improving health care quality for racial/ethnic minorities: a systematic review of the best evidence regarding provider and organization interventions, *BMC Public Health*, 6, 104.
38. BULLINGER, M., ALONSO, J., APOLONE, G. et al. (1998) Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment, *J Clin Epidemiol*, 51, 913-23.
39. ANDERSON, R. T., AARONSON, N. K., BULLINGER, M. & MCBEE, W. L. (1996) A review of the progress towards developing health-related quality-of-life instruments for international clinical studies and outcomes research, *Pharmacoeconomics*, 10, 336-55.
40. HUNT, S. M. (1986) Cross-cultural issues in the use of socio-medical indicators, *Health Policy*, 6, 149-58.
41. HERDMAN, M., FOX-RUSHBY, J. & BADIA, X. (1998) A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach, *Qual Life Res*, 7, 323-35.

42. WILD, D., EREMENCO, S., MEAR, I. et al. (2009) Multinational trials-recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report, *Value Health*, 12, 430-40.
43. ACQUADRO, C., CONWAY, K., HAREENDRAN, A. & AARONSON, N. (2008) Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials, *Value Health*, 11, 509-21.
44. ACQUADRO, C., BERZON, R., DUBOIS, D. et al. (2003) Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001, *Value Health*, 6, 522-31.
45. SKEVINGTON, S. M. (2002) Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. World Health Organisation Quality of Life Assessment, *Qual Life Res*, 11, 135-44.
46. GUYATT, G. H. (1993) The philosophy of health-related quality of life translation, *Qual Life Res*, 2, 461-5.
47. (2002) Assessing health status and quality-of-life instruments: attributes and review criteria, *Qual Life Res*, 11, 193-205.
48. LOHR, K. N., AARONSON, N. K., ALONSO, J. et al. (1996) Evaluating quality-of-life and health status instruments: development of scientific review criteria, *Clin Ther*, 18, 979-92.
49. BOERS, M., BROOKS, P., STRAND, C. V. & TUGWELL, P. (1998) The OMERACT filter for Outcome Measures in Rheumatology, *J Rheumatol*, 25, 198-9.
50. AKAI, M., DOI, T., FUJINO, K. et al. (2005) An outcome measure for Japanese people with knee osteoarthritis, *J Rheumatol*, 32, 1524-32.
51. MOKKINK, L. B., TERWEE, C. B., KNOL, D. L. et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content, *BMC Med Res Methodol*, 10, 22.
52. MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L. et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study, *Qual Life Res*, 19, 539-49.
53. MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L. et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes, *J Clin Epidemiol*, 63, 737-45.
54. FERMANIAN, J. (2005) [Validation of assessment scales in physical medicine and rehabilitation: how are psychometric properties determined?], *Ann Readapt Med Phys*, 48, 281-7.
55. CRONBACH, L. (1951) Coefficient alpha and the internal structure of tests, *Psychometrika*, 297-333.
56. BLAND, J. M. & ALTMAN, D. G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet*, 1, 307-10.
57. RAVAUD, P., GIRAudeau, B., AULELEY, G. R. et al. (1999) Assessing smallest detectable change over time in continuous structural outcome measures: application to radiological change in knee osteoarthritis, *J Clin Epidemiol*, 52, 1225-30.
58. WRIGHT, J. G. & YOUNG, N. L. (1997) A comparison of different indices of responsiveness, *J Clin Epidemiol*, 50, 239-46.
59. BELLAMY, N. (2005) Science of assessment, *Ann Rheum Dis*, 64 Suppl 2, ii42-5.
60. LEQUESNE, M. G. (1997) The algofunctional indices for hip and knee osteoarthritis, *J Rheumatol*, 24, 779-81.
61. JOHNSON, D. S. & SMITH, R. B. (2001) Outcome measurement in the ACL deficient knee--what's the score?, *Knee*, 8, 51-7.
62. DAVIES, A. P. (2002) Rating systems for total knee replacement, *Knee*, 9, 261-6.
63. PAXTON, E. W. & FITHIAN, D. C. (2005) Outcome instruments for patellofemoral arthroplasty, *Clin Orthop Relat Res*, 66-70.

64. WYLDE, V. & BLOM, A. W. (2009) Assessment of outcomes after hip arthroplasty, *Hip Int*, 19, 1-7.
65. WRIGHT, R. W. (2009) Knee injury outcomes measures, *J Am Acad Orthop Surg*, 17, 31-9.
66. SUN, Y., STURMER, T., GUNTHER, K. P. & BRENNER, H. (1997) Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee--a review of the literature, *Clin Rheumatol*, 16, 185-98.
67. VEENHOF, C., BIJLSMA, J. W., VAN DEN ENDE, C. H. et al. (2006) Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature, *Arthritis Rheum*, 55, 480-92.
68. CHAORY, K. & POIRAUDEAU, S. (2004) [Rating scores for ACL ligamentoplasty], *Ann Readapt Med Phys*, 47, 309-16.
69. BRYANT, D. M., SANDERS, D. W., COLES, C. P. et al. (2009) Selection of outcome measures for patients with hip fracture, *J Orthop Trauma*, 23, 434-41.
70. ETHGEN, O., BRUYERE, O., RICHY, F., DARDENNES, C. & REGINSTER, J. Y. (2004) Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature, *J Bone Joint Surg Am*, 86-A, 963-74.
71. EMERY, M. P., PERRIER, L. L. & ACQUADRO, C. (2005) Patient-reported outcome and quality of life instruments database (PROQOLID): frequently asked questions, *Health Qual Life Outcomes*, 3, 12.
72. MARSHALL, J. L., FETTO, J. F. & BOTERO, P. M. (1977) Knee ligament injuries: a standardized evaluation method, *Clin Orthop Relat Res*, 115-29.
73. POLLARD, B., DIXON, D., DIEPPE, P. & JOHNSTON, M. (2009) Measuring the ICF components of impairment, activity limitation and participation restriction: an item analysis using classical test theory and item response theory, *Health Qual Life Outcomes*, 7, 41.
74. BRAZIER, J. E., HARPER, R., JONES, N. M. et al. (1992) Validating the SF-36 health survey questionnaire: new outcome measure for primary care, *Bmj*, 305, 160-4.
75. HURST, N. P., KIND, P., RUTA, D., HUNTER, M. & STUBBINGS, A. (1997) Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D), *Br J Rheumatol*, 36, 551-9.
76. MEENAN, R. F., GERTMAN, P. M. & MASON, J. H. (1980) Measuring health status in arthritis. The arthritis impact measurement scales, *Arthritis Rheum*, 23, 146-52.
77. TEGNER, Y. & LYSHOLM, J. (1985) Rating systems in the evaluation of knee ligament injuries, *Clin Orthop Relat Res*, 43-9.
78. WILLIAMS, G. N., TAYLOR, D. C., GANGEL, T. J., UHORCHAK, J. M. & ARCIERO, R. A. (2000) Comparison of the single assessment numeric evaluation method and the Lysholm score, *Clin Orthop Relat Res*, 184-92.
79. ORNETTI, P., PERRUCCIO, A. V., ROOS, E. M. et al. (2009) Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS), *Osteoarthritis Cartilage*, 17, 1604-8.
80. ORNETTI, P., PARRATTE, S., GOSSEC, L. et al. Cross-cultural adaptation and validation of the French version of the Hip disability and Osteoarthritis Outcome Score (HOOS) in hip osteoarthritis patients, *Osteoarthritis Cartilage*, 18, 522-9.
81. NUNALLY, J. C. (1978) *Psychometric theory* (New-York, McGraw-Hill).
82. XIE, F., LI, S. C., ROOS, E. M. et al. (2006) Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore, *Osteoarthritis Cartilage*, 14, 1098-103.
83. COHEN, J. (1988) *Statistical power analysis for the behavioral sciences* (Hillsdale, NJ, Lawrence Erlbaum).
84. MARX, R. G., STUMP, T. J., JONES, E. C., WICKIEWICZ, T. L. & WARREN, R. F. (2001) Development and evaluation of an activity rating scale for disorders of the knee, *Am J Sports Med*, 29, 213-8.
85. JOHANSON, N. A., LIANG, M. H., DALTRY, L., RUDICEL, S. & RICHMOND, J. (2004) American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change, *J Bone Joint Surg Am*, 86-A, 902-9.

86. MARX, R. G., JONES, E. C., ALLEN, A. A. et al. (2001) Reliability, validity, and responsiveness of four knee outcome scales for athletic patients, *J Bone Joint Surg Am*, 83-A, 1459-69.
87. MOHTADI, N. (1998) Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency, *Am J Sports Med*, 26, 350-9.
88. HARTWICK, M., MEEUWISSE, W., VANDERTUIN, J. & MAITLAND, M. (2003) Knee pain in the ACL-deficient osteoarthritic knee and its relationship to quality of life, *Physiother Res Int*, 8, 83-92.
89. RISBERG, M. A., HOLM, I., STEEN, H. & BEYNNON, B. D. (1999) Sensitivity to changes over time for the IKDC form, the Lysholm score, and the Cincinnati knee score. A prospective study of 120 ACL reconstructed patients with a 2-year follow-up, *Knee Surg Sports Traumatol Arthrosc*, 7, 152-9.
90. BARBER-WESTIN, S. D., NOYES, F. R. & MCCLOSKEY, J. W. (1999) Rigorous statistical reliability, validity, and responsiveness testing of the Cincinnati knee rating system in 350 subjects with uninjured, injured, or anterior cruciate ligament-reconstructed knees, *Am J Sports Med*, 27, 402-16.
91. NOYES, F. R., MOOAR, P. A., MATTHEWS, D. S. & BUTLER, D. L. (1983) The symptomatic anterior cruciate-deficient knee. Part I: the long-term functional disability in athletically active individuals, *J Bone Joint Surg Am*, 65, 154-62.
92. NOYES, F. R., BARBER, S. D. & MOOAR, L. A. (1989) A rationale for assessing sports activity levels and limitations in knee disorders, *Clin Orthop Relat Res*, 238-49.
93. LEIGH BROWN, A. P., KENNEDY, A. D. M., GRANT, A. M. et al. (1999) The development and validation of the Edinburgh Knee Function Scale: a simple tool for outcome measurement in non-surgical patients, *The Knee*, 6, 115-123.
94. FULKERSON, J. P. & SHEA, K. P. (1990) Disorders of patellofemoral alignment, *J Bone Joint Surg Am*, 72, 1424-9.
95. PAXTON, E. W., FITHIAN, D. C., STONE, M. L. & SILVA, P. (2003) The reliability and validity of knee-specific and general health instruments in assessing acute patellar dislocation outcomes, *Am J Sports Med*, 31, 487-92.
96. CHESWORTH, B. M., CULHAM, E., TATA, G. E. & PEAT, M. (1989) Validation of outcome measures in patients with patellofemoral syndrome, *J Orthop Sports Phys Ther*, 10, 302-8.
97. CROSSLEY, K. M., BENNELL, K. L., COWAN, S. M. & GREEN, S. (2004) Analysis of outcome measures for persons with patellofemoral pain: which are reliable and valid?, *Arch Phys Med Rehabil*, 85, 815-22.
98. FLANDRY, F., HUNT, J. P., TERRY, G. C. & HUGHSTON, J. C. (1991) Analysis of subjective knee complaints using visual analog scales, *Am J Sports Med*, 19, 112-8.
99. HUGHSTON, J. C. & BARRETT, G. R. (1983) Acute anteromedial rotatory instability. Long-term results of surgical repair, *J Bone Joint Surg Am*, 65, 145-53.
100. IRRGANG, J. J., ANDERSON, A. F., BOLAND, A. L. et al. (2001) Development and validation of the international knee documentation committee subjective knee form, *Am J Sports Med*, 29, 600-13.
101. PADUA, R., BONDI, R., CECCARELLI, E. et al. (2004) Italian version of the International Knee Documentation Committee Subjective Knee Form: cross-cultural adaptation and validation, *Arthroscopy*, 20, 819-23.
102. HAVERKAMP, D., SIEREVELT, I. N., BREUGEM, S. J. et al. (2006) Translation and validation of the Dutch version of the International Knee Documentation Committee Subjective Knee Form, *Am J Sports Med*, 34, 1680-4.
103. LERTWANICH, P., PRAPHRUETKIT, T., KEYURAPAN, E., LAMSAM, C. & KULTHANAN, T. (2008) Validity and reliability of Thai version of the International Knee Documentation Committee Subjective Knee Form, *J Med Assoc Thai*, 91, 1218-25.
104. METSAVAHT, L., LEPORACE, G., RIBERTO, M., SPOSITO, M. M. & BATISTA, L. A. Translation and Cross-Cultural Adaptation of the Brazilian Version of the International Knee Documentation Committee Subjective Knee Form: Validity and Reproducibility, *Am J Sports Med*.

105. IRRGANG, J. J., ANDERSON, A. F., BOLAND, A. L. et al. (2006) Responsiveness of the International Knee Documentation Committee Subjective Knee Form, *Am J Sports Med*, 34, 1567-73.
106. BIZZINI, M. & GORELICK, M. (2007) Development of a German version of the knee outcome survey for daily activities, *Arch Orthop Trauma Surg*, 127, 781-9.
107. IRRGANG, J. J., SNYDER-MACKLER, L., WAINNER, R. S., FU, F. H. & HARNER, C. D. (1998) Development of a patient-reported measure of function of the knee, *J Bone Joint Surg Am*, 80, 1132-45.
108. GONCALVES, R. S., CABRI, J. & PINHEIRO, J. P. (2008) Cross-cultural adaptation and validation of the Portuguese version of the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS), *Clin Rheumatol*, 27, 1445-9.
109. EVCIK, D., AY, S., EGE, A., TUREL, A. & KAVUNCU, V. (2009) Adaptation and validation of Turkish version of the Knee Outcome Survey-Activities for Daily Living Scale, *Clin Orthop Relat Res*, 467, 2077-82.
110. GARRATT, A. M., BREALEY, S., ROBLING, M. et al. (2008) Development of the knee quality of life (KQoL-26) 26-item questionnaire: data quality, reliability, validity and responsiveness, *Health Qual Life Outcomes*, 6, 48.
111. LEQUESNE, M. G., MERY, C., SAMSON, M. & GERARD, P. (1987) Indexes of severity for osteoarthritis of the hip and knee. Validation--value in comparison with other assessment tests, *Scand J Rheumatol Suppl*, 65, 85-9.
112. LEQUESNE, M. G. & SAMSON, M. (1991) Indices of severity in osteoarthritis for weight bearing joints, *J Rheumatol Suppl*, 27, 16-8.
113. STUCKI, G., SANGHA, O., STUCKI, S. et al. (1998) Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis, *Osteoarthritis Cartilage*, 6, 79-86.
114. THEILER, R., SANGHA, O., SCHAEREN, S. et al. (1999) Superior responsiveness of the pain and function sections of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) as compared to the Lequesne-Algofunctional Index in patients with osteoarthritis of the lower extremities, *Osteoarthritis Cartilage*, 7, 515-9.
115. STUCKI, G., MEIER, D., STUCKI, S. et al. (1996) [Evaluation of a German questionnaire version of the Lequesne cox- and gonarthrosis indices], *Z Rheumatol*, 55, 50-7.
116. LUDWIG, F. J., MELZER, C., GRIMMIG, H. & DAALMANN, H. H. (2002) [Cross cultural adaptation of the lequesne algofunctional indices for german speaking patients with osteoarthritis of the hip and the knee], *Rehabilitation (Stuttg)*, 41, 249-57.
117. BAE, S. C., LEE, H. S., YUN, H. R. et al. (2001) Cross-cultural adaptation and validation of Korean Western Ontario and McMaster Universities (WOMAC) and Lequesne osteoarthritis indices for clinical research, *Osteoarthritis Cartilage*, 9, 746-50.
118. FAUCHER, M., POIRAUDEAU, S., LEFEVRE-COLAU, M. M. et al. (2003) Assessment of the test-retest reliability and construct validity of a modified Lequesne index in knee osteoarthritis, *Joint Bone Spine*, 70, 521-5.
119. GUERMAZI, M., MEZGANNI, M., YAHIA, M. et al. (2004) [Translation and construct validity of the Lequesne index for Arab speaking North African patients with osteoarthritis of the knee], *Ann Readapt Med Phys*, 47, 195-203.
120. XIE, F., THUMBOO, J., LO, N. N. et al. (2007) Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Lequesne Algofunctional Index of knee in Asians with knee osteoarthritis in Singapore, *Osteoarthritis Cartilage*, 15, 19-26.
121. VILLANUEVA, I., DEL MAR GUZMAN, M., JAVIER TOYOS, F., ARIZA-ARIZA, R. & NAVARRO, F. (2004) Relative efficiency and validity properties of a visual analogue vs a categorical scaled version of the Western Ontario and McMaster Universities Osteoarthritis (WOMAC) Index: Spanish versions, *Osteoarthritis Cartilage*, 12, 225-31.
122. KOCHER, M. S., STEADMAN, J. R., BRIGGS, K. K., STERETT, W. I. & HAWKINS, R. J. (2004) Reliability, validity, and responsiveness of the Lysholm knee scale for various chondral disorders of the knee, *J Bone Joint Surg Am*, 86-A, 1139-45.

123. BRIGGS, K. K., KOCHER, M. S., RODKEY, W. G. & STEADMAN, J. R. (2006) Reliability, validity, and responsiveness of the Lysholm knee score and Tegner activity scale for patients with meniscal injury of the knee, *J Bone Joint Surg Am*, 88, 698-705.
124. HEINTJES, E. M., BIERMA-ZEINSTR, S. M., BERGER, M. Y. & KOES, B. W. (2008) Lysholm scale and WOMAC index were responsive in prospective cohort of young general practice patients, *J Clin Epidemiol*, 61, 481-8.
125. SMITH, H. J., RICHARDSON, J. B. & TENNANT, A. (2009) Modification and validation of the Lysholm Knee Scale to assess articular cartilage damage, *Osteoarthritis Cartilage*, 17, 53-8.
126. BRIGGS, K. K., LYSHOLM, J., TEGNER, Y. et al. (2009) The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee: 25 years later, *Am J Sports Med*, 37, 890-7.
127. BENGTSSON, J., MOLLBORG, J. & WERNER, S. (1996) A study for testing the sensitivity and reliability of the Lysholm knee scoring scale, *Knee Surg Sports Traumatol Arthrosc*, 4, 27-31.
128. LYSHOLM, J. & GILLQUIST, J. (1982) Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale, *Am J Sports Med*, 10, 150-4.
129. LIOW, R. Y., WALKER, K., WAJID, M. A., BEDI, G. & LENNOX, C. M. (2003) Functional rating for knee arthroplasty: comparison of three scoring systems, *Orthopedics*, 26, 143-9.
130. DAWSON, J., FITZPATRICK, R., MURRAY, D. & CARR, A. (1998) Questionnaire on the perceptions of patients about total knee replacement, *J Bone Joint Surg Br*, 80, 63-9.
131. DUNBAR, M. J., ROBERTSSON, O., RYD, L. & LIDGREN, L. (2000) Translation and validation of the Oxford-12 item knee score for use in Sweden, *Acta Orthop Scand*, 71, 268-74.
132. CHAROENCHOLVANICH, K. & PONGCHAROEN, B. (2005) Oxford knee score and SF-36: translation & reliability for use with total knee arthroscopy patients in Thailand, *J Med Assoc Thai*, 88, 1194-202.
133. XIE, F., LI, S. C., LO, N. N. et al. (2007) Cross-cultural adaptation and validation of Singapore English and Chinese Versions of the Oxford Knee Score (OKS) in knee osteoarthritis patients undergoing total knee replacement, *Osteoarthritis Cartilage*, 15, 1019-24.
134. CONAGHAN, P. G., EMERTON, M. & TENNANT, A. (2007) Internal construct validity of the Oxford Knee Scale: evidence from Rasch measurement, *Arthritis Rheum*, 57, 1363-7.
135. NAAL, F. D., IMPELLIZZERI, F. M., SIEVERDING, M. et al. (2009) The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee, *Osteoarthritis Cartilage*, 17, 49-52.
136. LAPRADE, J. A. & CULHAM, E. G. (2002) A self-administered pain severity scale for patellofemoral pain syndrome, *Clin Rehabil*, 16, 780-8.
137. ZUCKERMAN, J. D., KOVAL, K. J., AHARONOFF, G. B., HIEBERT, R. & SKOVRON, M. L. (2000) A functional recovery score for elderly hip fracture patients: I. Development, *J Orthop Trauma*, 14, 20-5.
138. ZUCKERMAN, J. D., KOVAL, K. J., AHARONOFF, G. B. & SKOVRON, M. L. (2000) A functional recovery score for elderly hip fracture patients: II. Validity and reliability, *J Orthop Trauma*, 14, 26-30.
139. MARTIN, R. L. & PHILIPPON, M. J. (2007) Evidence of validity for the hip outcome score in hip arthroscopy, *Arthroscopy*, 23, 822-6.
140. MARTIN, R. L., KELLY, B. T. & PHILIPPON, M. J. (2006) Evidence of validity for the hip outcome score, *Arthroscopy*, 22, 1304-11.
141. MARTIN, R. L. & PHILIPPON, M. J. (2008) Evidence of reliability and responsiveness for the hip outcome score, *Arthroscopy*, 24, 676-82.
142. SEKIYA, J. K., MARTIN, R. L. & LESNIAK, B. P. (2009) Arthroscopic repair of delaminated acetabular articular cartilage in femoroacetabular impingement, *Orthopedics*, 32.
143. JOHANSON, N. A., CHARLSON, M. E., SZATROWSKI, T. P. & RANAWAT, C. S. (1992) A self-administered hip-rating questionnaire for the assessment of outcome after total hip replacement, *J Bone Joint Surg Am*, 74, 587-97.
144. ZIMMERMAN, S., HAWKES, W. G., HEBEL, J. R. et al. (2006) The Lower Extremity Gain Scale: a performance-based measure to assess recovery after hip fracture, *Arch Phys Med Rehabil*, 87, 430-6.

145. HAWKES, W. G., WILLIAMS, G. R., ZIMMERMAN, S. et al. (2004) A clinically meaningful difference was generated for a performance measure of recovery from hip fracture, *J Clin Epidemiol*, 57, 1019-24.
146. JAGLAL, S., LAKHANI, Z. & SCHATZKER, J. (2000) Reliability, validity, and responsiveness of the lower extremity measure for patients with a hip fracture, *J Bone Joint Surg Am*, 82-A, 955-62.
147. LAUDE, F., SARIALI, E. & NOGIER, A. (2009) Femoroacetabular impingement treatment using arthroscopy and anterior approach, *Clin Orthop Relat Res*, 467, 747-52.
148. STAHELIN, L., STAHELIN, T., JOLLES, B. M. & HERZOG, R. F. (2008) Arthroscopic offset restoration in femoroacetabular cam impingement: accuracy and early clinical outcome, *Arthroscopy*, 24, 51-57 e1.
149. CHRISTENSEN, C. P., ALTHAUSEN, P. L., MITTLEMAN, M. A., LEE, J. A. & MCCARTHY, J. C. (2003) The nonarthritic hip score: reliable and validated, *Clin Orthop Relat Res*, 75-83.
150. DAWSON, J., FITZPATRICK, R., CARR, A. & MURRAY, D. (1996) Questionnaire on the perceptions of patients about total hip replacement, *J Bone Joint Surg Br*, 78, 185-90.
151. DAWSON, J., FITZPATRICK, R., MURRAY, D. & CARR, A. (1996) Comparison of measures to assess outcomes in total hip replacement surgery, *Qual Health Care*, 5, 81-8.
152. DAWSON, J., FITZPATRICK, R., FROST, S. et al. (2001) Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement, *J Bone Joint Surg Br*, 83, 1125-9.
153. UESUGI, Y., MAKIMOTO, K., FUJITA, K. et al. (2009) Validity and responsiveness of the Oxford hip score in a prospective study with Japanese total hip arthroplasty patients, *J Orthop Sci*, 14, 35-9.
154. DELAUNAY, C., EPINETTE, J. A., DAWSON, J., MURRAY, D. & JOLLES, B. M. (2009) Cross-cultural adaptations of the Oxford-12 HIP score to the French speaking population, *Orthop Traumatol Surg Res*, 95, 89-99.
155. NAAL, F. D., SIEVERDING, M., IMPELLIZZERI, F. M. et al. (2009) Reliability and validity of the cross-culturally adapted German Oxford hip score, *Clin Orthop Relat Res*, 467, 952-7.
156. DE GROOT, I. B., REIJMAN, M., TERWEE, C. B. et al. (2007) Validation of the Dutch version of the Hip disability and Osteoarthritis Outcome Score, *Osteoarthritis Cartilage*, 15, 104-9.
157. WRIGHT, J. G. & YOUNG, N. L. (1997) The patient-specific index: asking patients what they want, *J Bone Joint Surg Am*, 79, 974-83.
158. WRIGHT, J. G., YOUNG, N. L. & WADDELL, J. P. (2000) The reliability and validity of the self-reported patient-specific index for total hip arthroplasty, *J Bone Joint Surg Am*, 82, 829-37.
159. PARKER, M. J. & PALMER, C. R. (1993) A new mobility score for predicting mortality after hip fracture, *J Bone Joint Surg Br*, 75, 797-8.
160. LIANG, M. H., KATZ, J. N., PHILLIPS, C., SLEDGE, C. & CATS-BARIL, W. (1991) The total hip arthroplasty outcome evaluation form of the American Academy of Orthopaedic Surgeons. Results of a nominal group process. The American Academy of Orthopaedic Surgeons Task Force on Outcome Studies, *J Bone Joint Surg Am*, 73, 639-46.
161. KATZ, J. N., PHILLIPS, C. B., POSS, R. et al. (1995) The validity and reliability of a Total Hip Arthroplasty Outcome Evaluation Questionnaire, *J Bone Joint Surg Am*, 77, 1528-34.
162. KIRMIT, L., KARATOSUN, V., UNVER, B. et al. (2005) The reliability of hip scoring systems for total hip arthroplasty candidates: assessment by physical therapists, *Clin Rehabil*, 19, 659-61.
163. TALBOT, S., HOOPER, G., STOKES, A. & ZORDAN, R. Use of a new high-activity arthroplasty score to assess function of young patients with total hip or knee arthroplasty, *J Arthroplasty*, 25, 268-73.
164. GONCALVES, R. S., CABRI, J., PINHEIRO, J. P., FERREIRA, P. L. & GIL, J. Cross-cultural adaptation and validation of the Portuguese version of the Intermittent and Constant OsteoArthritis Pain (ICOAP) measure for the knee, *Osteoarthritis Cartilage*.
165. DAVIS, A. M., LOHMANDER, L. S., WONG, R., VENKATARAMANAN, V. & HAWKER, G. A. Evaluating the responsiveness of the ICOAP following hip or knee replacement, *Osteoarthritis Cartilage*.

166. MAILLEFERT, J. F., KLOPPENBURG, M., FERNANDES, L. et al. (2009) Multi-language translation and cross-cultural adaptation of the OARSI/OMERACT measure of intermittent and constant osteoarthritis pain (ICOAP), *Osteoarthritis Cartilage*, 17, 1293-6.
167. HAWKER, G. A., DAVIS, A. M., FRENCH, M. R. et al. (2008) Development and preliminary psychometric testing of a new OA pain measure--an OARSI/OMERACT initiative, *Osteoarthritis Cartilage*, 16, 409-14.
168. SALEH, K. J., MULHALL, K. J., BERSHADSKY, B. et al. (2005) Development and validation of a lower-extremity activity scale. Use for patients treated with revision total knee arthroplasty, *J Bone Joint Surg Am*, 87, 1985-94.
169. GHOMRAWI, H. M., KANE, R. L., EBERLY, L. E., BERSHADSKY, B. & SALEH, K. J. (2009) Patterns of functional improvement after revision knee arthroplasty, *J Bone Joint Surg Am*, 91, 2838-45.
170. BINKLEY, J. M., STRATFORD, P. W., LOTT, S. A. & RIDDLE, D. L. (1999) The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network, *Phys Ther*, 79, 371-83.
171. STRATFORD, P. W., KENNEDY, D. M. & HANNA, S. E. (2004) Condition-specific Western Ontario McMaster Osteoarthritis Index was not superior to region-specific Lower Extremity Functional Scale at detecting change, *J Clin Epidemiol*, 57, 1025-32.
172. WATSON, C. J., PROPPS, M., RATNER, J. et al. (2005) Reliability and responsiveness of the lower extremity functional scale and the anterior knee pain scale in patients with anterior knee pain, *J Orthop Sports Phys Ther*, 35, 136-46.
173. KENNEDY, D. M., STRATFORD, P. W., RIDDLE, D. L., HANNA, S. E. & GOLLISH, J. D. (2008) Assessing recovery and establishing prognosis following total knee arthroplasty, *Phys Ther*, 88, 22-32.
174. STENGEL, D., KLUFMOLLER, F., RADEMACHER, G. et al. (2009) Functional outcomes and health-related quality of life after robot-assisted anterior cruciate ligament reconstruction with patellar tendon grafts, *Knee Surg Sports Traumatol Arthrosc*, 17, 446-55.
175. PUA, Y. H., COWAN, S. M., WRIGLEY, T. V. & BENNELL, K. L. (2009) The Lower Extremity Functional Scale could be an alternative to the Western Ontario and McMaster Universities Osteoarthritis Index physical function scale, *J Clin Epidemiol*, 62, 1103-11.
176. TOYE, F., BARLOW, J., WRIGHT, C. & LAMB, S. E. (2007) A validation study of the New Zealand score for hip and knee surgery, *Clin Orthop Relat Res*, 464, 190-5.
177. HADORN, D. C. & HOLMES, A. C. (1997) The New Zealand priority criteria project. Part 1: Overview, *Bmj*, 314, 131-4.
178. DERRETT, S., DEVLIN, N., HANSEN, P. & HERBISON, P. (2003) Prioritizing patients for elective surgery: a prospective study of clinical priority assessment criteria in New Zealand, *Int J Technol Assess Health Care*, 19, 91-105.
179. RAT, A. C., COSTE, J., POUCHOT, J. et al. (2005) OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis, *J Clin Epidemiol*, 58, 47-55.
180. RAT, A. C., POUCHOT, J., COSTE, J. et al. (2006) Development and testing of a specific quality-of-life questionnaire for knee and hip osteoarthritis: OAKHQOL (OsteoArthritis of Knee Hip Quality Of Life), *Joint Bone Spine*, 73, 697-704.
181. WHITEHOUSE, S. L., LINGARD, E. A., KATZ, J. N. & LEARMONTH, I. D. (2003) Development and testing of a reduced WOMAC function scale, *J Bone Joint Surg Br*, 85, 706-11.
182. WHITEHOUSE, S. L., CRAWFORD, R. W. & LEARMONTH, I. D. (2008) Validation for the reduced Western Ontario and McMaster Universities Osteoarthritis Index function scale, *J Orthop Surg (Hong Kong)*, 16, 50-3.
183. BREMANDER, A. B., PETERSSON, I. F. & ROOS, E. M. (2003) Validation of the Rheumatoid and Arthritis Outcome Score (RAOS) for the lower extremity, *Health Qual Life Outcomes*, 1, 55.
184. KARATEPE, A. G., GUNAYDIN, R., TURKMEN, G., KAYA, T. & OZBEK, G. (2009) The validity and reliability of the Turkish version of the Rheumatoid and Arthritis Outcome Score (RAOS) in patients with rheumatoid arthritis, *Clin Rheumatol*, 28, 47-51.

185. DUVAL, A., MAILLEFERT, J. F., GOSSEC, L. et al. (2010) Cross-cultural adaptation and validation of the french version of the Rheumatoid And Arthritis Outcome Score (RAOS), *Clin Exp Rheumatol*, in press.
186. GUERMAZI, M., POIRAUDEAU, S., YAHIA, M. et al. (2004) Translation, adaptation and validation of the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) for an Arab population: the Sfax modified WOMAC, *Osteoarthritis Cartilage*, 12, 459-68.
187. BELLAMY, N., CAMPBELL, J., STEVENS, J. et al. (1997) Validation study of a computerized version of the Western Ontario and McMaster Universities VA3.0 Osteoarthritis Index, *J Rheumatol*, 24, 2413-5.
188. WIGLER, I., NEUMANN, L. & YARON, M. (1999) Validation study of a Hebrew version of WOMAC in patients with osteoarthritis of the knee, *Clin Rheumatol*, 18, 402-5.
189. THUMBOO, J., CHEW, L. H. & SOH, C. H. (2001) Validation of the Western Ontario and McMaster University osteoarthritis index in Asians with osteoarthritis in Singapore, *Osteoarthritis Cartilage*, 9, 440-6.
190. MCCONNELL, S., KOLOPACK, P. & DAVIS, A. M. (2001) The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties, *Arthritis Rheum*, 45, 453-61.
191. ESCOBAR, A., QUINTANA, J. M., BILBAO, A., AZKARATE, J. & GUENAGA, J. I. (2002) Validation of the Spanish version of the WOMAC questionnaire for patients with hip or knee osteoarthritis. Western Ontario and McMaster Universities Osteoarthritis Index, *Clin Rheumatol*, 21, 466-71.
192. HASHIMOTO, H., HANYU, T., SLEDGE, C. B. & LINGARD, E. A. (2003) Validation of a Japanese patient-derived outcome scale for assessing total knee arthroplasty: comparison with Western Ontario and McMaster Universities osteoarthritis index (WOMAC), *J Orthop Sci*, 8, 288-93.
193. ROORDA, L. D., JONES, C. A., WALTZ, M. et al. (2004) Satisfactory cross cultural equivalence of the Dutch WOMAC in patients with hip osteoarthritis waiting for arthroplasty, *Ann Rheum Dis*, 63, 36-42.
194. TUZUN, E. H., EKER, L., AYTAR, A., DASKAPAN, A. & BAYRAMOGLU, M. (2005) Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index, *Osteoarthritis Cartilage*, 13, 28-33.
195. ANGST, F., EWERT, T., LEHMANN, S., AESCHLIMANN, A. & STUCKI, G. (2005) The factor subdimensions of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) help to specify hip and knee osteoarthritis. a prospective evaluation and validation study, *J Rheumatol*, 32, 1324-30.
196. KUPTNIRATSAIKUL, V. & RATTANACHAIYANONT, M. (2007) Validation of a modified Thai version of the Western Ontario and McMaster (WOMAC) osteoarthritis index for knee osteoarthritis, *Clin Rheumatol*, 26, 1641-5.
197. FAIK, A., BENBOUAZZA, K., AMINE, B. et al. (2008) Translation and validation of Moroccan Western Ontario and McMaster Universities (WOMAC) osteoarthritis index in knee osteoarthritis, *Rheumatol Int*, 28, 677-83.
198. XIE, F., LI, S. C., GOEREE, R. et al. (2008) Validation of Chinese Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement, *Qual Life Res*, 17, 595-601.
199. SOININEN, J. V., PAAVOLAINEN, P. O., GRONBLAD, M. A. & KAAPA, E. H. (2008) Validation study of a Finnish version of the Western Ontario and McMasters University osteoarthritis index, *Hip Int*, 18, 108-11.
200. BRUCE, B. & FRIES, J. (2004) Longitudinal comparison of the Health Assessment Questionnaire (HAQ) and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), *Arthritis Rheum*, 51, 730-7.
201. SODERMAN, P. & MALCHAU, H. (2000) Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP), *Acta Orthop Scand*, 71, 39-46.
202. SODERMAN, P. & MALCHAU, H. (2001) Is the Harris hip score system useful to study the outcome of total hip replacement?, *Clin Orthop Relat Res*, 189-97.

203. SODERMAN, P., MALCHAU, H. & HERBERTS, P. (2001) Outcome of total hip replacement: a comparison of different measurement methods, *Clin Orthop Relat Res*, 163-72.
204. BARON, G., TUBACH, F., RAVAUD, P., LOGEART, I. & DOUGADOS, M. (2007) Validation of a short form of the Western Ontario and McMaster Universities Osteoarthritis Index function subscale in hip and knee osteoarthritis, *Arthritis Rheum*, 57, 633-8.
205. TUBACH, F., BARON, G., FALISSARD, B. et al. (2005) Using patients' and rheumatologists' opinions to specify a short form of the WOMAC function subscale, *Ann Rheum Dis*, 64, 75-9.
206. WARD, M. M. (2004) Outcome measurement: health status and quality of life, *Curr Opin Rheumatol*, 16, 96-101.
207. BOMBARDIER, C., MELFI, C. A., PAUL, J. et al. (1995) Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery, *Med Care*, 33, AS131-44.
208. BEDSON, J. & CROFT, P. R. (2008) The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature, *BMC Musculoskelet Disord*, 9, 116.
209. KEEN, H. I., BINGHAM, C. O., 3RD, BRADLEY, L. A. et al. (2009) Assessing single joints in arthritis clinical trials, *J Rheumatol*, 36, 2092-6.
210. BARBER-WESTIN, S. D. & NOYES, F. R. (1999) Assessment of sports participation levels following knee injuries, *Sports Med*, 28, 1-10.
211. ANDREW, N. E., GABBE, B. J., WOLFE, R. & CAMERON, P. A. Evaluation of instruments for measuring the burden of sport and active recreation injury, *Sports Med*, 40, 141-61.
212. ACKERMAN, I. N., GRAVES, S. E., BENNELL, K. L. & OSBORNE, R. H. (2006) Evaluating quality of life in hip and knee replacement: Psychometric properties of the World Health Organization Quality of Life short version instrument, *Arthritis Rheum*, 55, 583-90.
213. BOUTRON, I., RANNOU, F., JARDINAUD-LOPEZ, M. et al. (2008) Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year, *Osteoarthritis Cartilage*, 16, 1024-31.
214. JENKINSON, C., COULTER, A. & WRIGHT, L. (1993) Short form 36 (SF36) health survey questionnaire: normative data for adults of working age, *Bmj*, 306, 1437-40.
215. TUBACH, F., WELLS, G. A., RAVAUD, P. & DOUGADOS, M. (2005) Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues, *J Rheumatol*, 32, 2025-9.
216. ROOS, E. M., OSTENBERG, A., ROOS, H., EKDAHL, C. & LOHMANDER, L. S. (2001) Long-term outcome of meniscectomy: symptoms, function, and performance tests in patients with or without radiographic osteoarthritis compared to matched controls, *Osteoarthritis Cartilage*, 9, 316-24.
217. <http://www.koos.nu/>.
218. ROOS, E. M. & LOHMANDER, L. S. (2003) The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis, *Health Qual Life Outcomes*, 1, 64.
219. NILSDOTTER, A. K., PETERSSON, I. F., ROOS, E. M. & LOHMANDER, L. S. (2003) Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study, *Ann Rheum Dis*, 62, 923-30.
220. PERRUCCIO, A. V., STEFAN LOHMANDER, L., CANIZARES, M. et al. (2008) The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative, *Osteoarthritis Cartilage*, 16, 542-50.
221. DAVIS, A. M., PERRUCCIO, A. V., CANIZARES, M. et al. (2008) The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative, *Osteoarthritis Cartilage*, 16, 551-9.
222. GOSSEC, L., HAWKER, G., DAVIS, A. M. et al. (2007) OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis, *J Rheumatol*, 34, 1432-5.
223. TENNANT, A. & CONAGHAN, P. G. (2007) The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper?, *Arthritis Rheum*, 57, 1358-62.

224. DAVIS, A. M., PERRUCCIO, A. V., CANIZARES, M. et al. (2009) Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis, *Osteoarthritis Cartilage*, 17, 843-7.
225. SALAFFI, F., CAROTTI, M. & GRASSI, W. (2005) Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments, *Clin Rheumatol*, 24, 29-37.
226. SHIELDS, R. K., ENLOE, L. J. & LEO, K. C. (1999) Health related quality of life in patients with total hip or knee replacement, *Arch Phys Med Rehabil*, 80, 572-9.
227. MALY, M. R., COSTIGAN, P. A. & OLNEY, S. J. (2006) Determinants of self-report outcome measures in people with knee osteoarthritis, *Arch Phys Med Rehabil*, 87, 96-104.
228. ROOS, E. M. & TOKSVIG-LARSEN, S. (2003) Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement, *Health Qual Life Outcomes*, 1, 17.
229. ROOS, E. M. & LOHMANDER, L. S. (2003) The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis, *Health Qual Life Outcomes*, 1, 64.
230. HUCH, K., MULLER, K. A., STURMER, T. et al. (2005) Sports activities 5 years after total knee or hip arthroplasty: the Ulm Osteoarthritis Study, *Ann Rheum Dis*, 64, 1715-20.
231. ROOS, E. M. & TOKSVIG-LARSEN, S. (2003) Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement, *Health Qual Life Outcomes*, 1, 17.
232. KESSLER, S., LANG, S., PUHL, W. & STOVE, J. (2003) [The Knee Injury and Osteoarthritis Outcome Score--a multifunctional questionnaire to measure outcome in knee arthroplasty], *Z Orthop Ihre Grenzgeb*, 141, 277-82.
233. ALTMAN, R., ASCH, E., BLOCH, D. et al. (1986) Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association, *Arthritis Rheum*, 29, 1039-49.
234. XIE, F., LI, S. C., ROOS, E. M. et al. (2006) Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore, *Osteoarthritis Cartilage*.
235. PARADOWSKI, P. T., ENGLUND, M., LOHMANDER, L. S. & ROOS, E. M. (2005) The effect of patient characteristics on variability in pain and function over two years in early knee osteoarthritis, *Health Qual Life Outcomes*, 3, 59.
236. ALTMAN, R., ALARCON, G., APPELROUTH, D. et al. (1991) The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip, *Arthritis Rheum*, 34, 505-14.
237. GEORGE, D. & MALLERY, P. (2003) *SPSS for Windows Step by Step: A Simple Guide and Reference 11.0 update* (Boston, Allyn & Bacon).
238. LANDIS, J. R. & KOCH, G. G. (1977) The measurement of observer agreement for categorical data, *Biometrics*, 33, 159-74.
239. FLEISS, J. L. & SHROUT, P. (1977) The effects of measurement errors on some multivariate procedures, *Am J Public Health*, 67, 1188-91.
240. ORNETTI, P., PARRATTE, S., GOSSEC, L. et al. (2008) Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients, *Osteoarthritis Cartilage*, 16, 423-8.
241. MCHORNEY, C. & TARLOV, A. (1995) Individual-patient monitoring in clinical practice: are available health status surveys adequate?, *Qual Life Res*, 4, 293-307.
242. RICHETTE, P., RAVAUD, P., CONROZIER, T. et al. (2009) Effect of hyaluronic acid in symptomatic hip osteoarthritis: a multicenter, randomized, placebo-controlled trial, *Arthritis Rheum*, 60, 824-30.
243. QVISTGAARD, E., CHRISTENSEN, R., TORP-PEDERSEN, S. & BLIDDAL, H. (2006) Intra-articular treatment of hip osteoarthritis: a randomized trial of hyaluronic acid, corticosteroid, and isotonic saline, *Osteoarthritis Cartilage*, 14, 163-70.

244. VAN DEN BEKEROM, M. P., LAMME, B., SERMON, A. & MULIER, M. (2008) What is the evidence for viscosupplementation in the treatment of patients with hip osteoarthritis? Systematic review of the literature, *Arch Orthop Trauma Surg*, 128, 815-23.
245. ZHANG, W., ROBERTSON, J., JONES, A. C., DIEPPE, P. A. & DOHERTY, M. (2008) The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials, *Ann Rheum Dis*, 67, 1716-23.
246. LAMBERT, R. G., HUTCHINGS, E. J., GRACE, M. G. et al. (2007) Steroid injection for osteoarthritis of the hip: a randomized, double-blind, placebo-controlled trial, *Arthritis Rheum*, 56, 2278-87.
247. FERNANDEZ LOPEZ, J. C. & RUANO-RAVINA, A. (2006) Efficacy and safety of intraarticular hyaluronic acid in the treatment of hip osteoarthritis: a systematic review, *Osteoarthritis Cartilage*, 14, 1306-11.
248. BERG, P. & OLSSON, U. (2004) Intra-articular injection of non-animal stabilised hyaluronic acid (NASHA) for osteoarthritis of the hip: a pilot study, *Clin Exp Rheumatol*, 22, 300-6.
249. DE GROOT, I. B., FAVEJEE, M. M., REIJMAN, M., VERHAAR, J. A. & TERWEE, C. B. (2008) The Dutch version of the Knee Injury and Osteoarthritis Outcome Score: a validation study, *Health Qual Life Outcomes*, 6, 16.
250. SALAVATI, M., MAZAHARI, M., NEGAHBAN, H. et al. (2008) Validation of a Persian-version of Knee injury and Osteoarthritis Outcome Score (KOOS) in Iranians with knee injuries, *Osteoarthritis Cartilage*, 16, 1178-82.
251. HAMBLY, K. & GRIVA, K. (2008) IKDC or KOOS? Which measures symptoms and disabilities most important to postoperative articular cartilage repair patients?, *Am J Sports Med*, 36, 1695-704.
252. GONCALVES, R. S., CABRI, J., PINHEIRO, J. P. & FERREIRA, P. L. (2009) Cross-cultural adaptation and validation of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score (KOOS), *Osteoarthritis Cartilage*, 17, 1156-62.
253. BEKKERS, J. E., DE WINDT, T. S., RAJMAKERS, N. J., DHERT, W. J. & SARIS, D. B. (2009) Validation of the Knee Injury and Osteoarthritis Outcome Score (KOOS) for the treatment of focal cartilage lesions, *Osteoarthritis Cartilage*, 17, 1434-9.
254. ROOS, E. M., ROOS, H. P. & LOHMANDER, L. S. (1999) WOMAC Osteoarthritis Index--additional dimensions for use in subjects with post-traumatic osteoarthritis of the knee. Western Ontario and MacMaster Universities, *Osteoarthritis Cartilage*, 7, 216-21.
255. ROOS, E. M., ROOS, H. P., RYD, L. & LOHMANDER, L. S. (2000) Substantial disability 3 months after arthroscopic partial meniscectomy: A prospective study of patient-relevant outcomes, *Arthroscopy*, 16, 619-26.
256. GONCALVES, R. S., CABRI, J., PINHEIRO, J. P., FERREIRA, P. L. & GIL, J. Reliability, validity and responsiveness of the Portuguese version of the Knee injury and Osteoarthritis Outcome Score--Physical Function Short-form (KOOS-PS), *Osteoarthritis Cartilage*, 18, 372-6.
257. CHEVALIER, X. Intraarticular treatments for osteoarthritis: new perspectives, *Curr Drug Targets*, 11, 546-60.
258. MEASE, P. J., HOBBS, K., CHALMERS, A. et al. (2009) Local delivery of a recombinant adenoassociated vector containing a tumour necrosis factor alpha antagonist gene in inflammatory arthritis: a phase 1 dose-escalation safety and tolerability study, *Ann Rheum Dis*, 68, 1247-54.
259. GOSSEC, L., DOUGADOS, M., RINCHEVAL, N. et al. (2009) Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative, *Ann Rheum Dis*, 68, 1680-5.
260. MAILLEFERT, J. F., GUEGUEN, A., NGUYEN, M. et al. (2002) A composite index for total hip arthroplasty in patients with hip osteoarthritis, *J Rheumatol*, 29, 347-52.

Annexes.

PUBLICATIONS ISSUES DU TRAVAIL DE THESE

- **Annexe 1.** Article KOOS

Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, Guillemin F, Maillefert JF. Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. *Osteoarthritis Cartilage*. 2008 Apr;16(4):423-8.

- **Annexe 2.** Article HOOS

Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, Guillemin F, Maillefert JF. Cross-cultural adaptation and validation of the French version of the Hip disability and Osteoarthritis Outcome Score (HOOS) in hip osteoarthritis patients. *Osteoarthritis Cartilage*. 2010 Apr;18(4):522-9.

- **Annexe 3.** Article KOOS-PS/HOOS-PS

Ornetti P, Perruccio AV, Roos EM, Lohmander LS, Davis AM, Maillefert JF. Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS). *Osteoarthritis Cartilage*. 2009 Dec;17(12):1604-8.

VERSIONS FRANÇAISES DU KOOS ET DU HOOS

(disponibles en format PDF sur internet à l'adresse suivante <http://www.koos.nu>)

- **Annexe 4.** Version française du KOOS

- **Annexe 5.** Version française du HOOS

Annexe 1. Article KOOS

Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients¹

P. Ornetti M.D.^{†‡§}, S. Parratte M.D.^{||}, L. Gossec M.D.^{¶#}, C. Tavernier M.D.^{†‡}, J.-N. Argenson M.D., Ph.D.^{||}, E. M. Roos P.T., Ph.D.^{†‡§§}, F. Guillemin M.D., Ph.D.^{†‡} and J. F. Maillefer M.D., Ph.D.^{†‡§*}

[†] Department of Rheumatology, Dijon University Hospital, F-21079, France

[‡] Université de Bourgogne, Faculty of Medicine, Dijon F-21078, France

[§] INSERM U887, Dijon F-21078, France

^{||} Department of Orthopedic Surgery, Hôpital Sainte Marguerite, Aix-Marseille University, Marseille, France

[¶] Rhumatologie B, Cochin Hospital, Paris, France

[#] Paris V René Descartes University, Paris, France

^{††} Department of Orthopedics, Clinical Sciences Lund, Lund University, Sweden

^{‡‡} EA4003, Nancy University, Clinical Epidemiology Department, INSERM CIE6, Marin Hospital, Nancy, France

^{§§} Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Denmark

Summary

Objective: To adapt the Knee injury and Osteoarthritis Outcome Score (KOOS) into French and to evaluate the psychometric properties of this new version.

Methods: The French version of the KOOS was developed according to cross-cultural guidelines by using the “translation-back translation” method to ensure content validity. KOOS data were then obtained in patients with symptomatic knee osteoarthritis (OA). The translated questionnaire was evaluated in two knee OA population groups, one with no indication for joint replacement (medicine), and the other waiting for joint replacement (surgery). The psychometric properties evaluated were feasibility: percentage of responses, floor and ceiling effects; construct validity: internal consistency using Cronbach’s alpha, correlations with osteoarthritis knee and hip quality of life domains using Spearman’s rank test, and known group comparison between medicine and surgery groups; reliability: intra-class correlation coefficient (ICC), Bland and Altman representation; responsiveness using data obtained prior to and 3 months after surgery: standardized response mean (SRM), and effect size.

Results: Thirty-seven patients were included in the medicine group (68% women, mean age = 70 ± 10 years) and 30 in the surgery group (73% women, mean age = 71 ± 10 years). The percentage of responses was excellent. Neither a floor nor a ceiling effect was observed, except for the sport and recreation subscale (20.6% of patients with the worst possible score in the medicine group, 40 and 0% in the surgery group prior to and after surgery, respectively). Results for internal consistency (Cronbach’s alpha ranging from 0.76 to 0.93), and convergent and divergent construct validity were satisfactory. The patients waiting for knee surgery presented with significantly lower scores in all KOOS domains. The reproducibility of measurements of all KOOS subscales was good to excellent, with ICC ranging from 0.755 to 0.914. The responsiveness was high, with SRM ranging from 0.89 to 1.93, and effect size from 1.31 to 2.8.

Conclusion: The French version of KOOS is a valid, reliable, and responsive instrument to capture specific aspects of functional disability affecting quality of life of knee OA patients.

© 2007 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Key words: Knee osteoarthritis, KOOS, Cross-cultural validation.

Osteoarthritis (OA) is a common, degenerative joint disease characterized by progressive destruction of cartilage, affecting to large extent weight-bearing joints, such as the knee.

The pain and disability associated with knee OA have a significant impact on the patients’ health-related quality of life (QOL)^{1,2}. Various instruments are available to assess knee in OA patients^{1,3–5}. In particular, the Western Ontario and McMaster Universities Index (WOMAC) is a validated and widely used disease-specific instrument, which assesses OA-induced pain, stiffness, and functional limitation⁶. The Knee injury and Osteoarthritis Outcome Score (KOOS) was developed as an extension of the WOMAC for young and/or active patients with knee OA or knee injury^{7,8}. The main difference between the WOMAC and the KOOS is the inclusion of the sport and recreation function and QOL domains. The

¹This study was supported by the Dijon University Hospital, France.

*Address correspondence and reprint requests to: Dr J. F. Maillefer, M.D., Ph.D., Department of Rheumatology, Hôpital Général, 3 rue du Faubourg Raines, 21000 Dijon, France. Tel: 33-3-80-29-37-45; Fax: 33-3-80-29-36-78; E-mail: jean-francis.maillefer@chu-dijon.fr

Received 16 December 2006; revision accepted 13 August 2007.

KOOS has been shown to be more sensitive and responsive than WOMAC in younger or more active patient^{8–10}. Several studies suggest that the current aging population, which includes those with knee OA awaiting joint replacement, is physically more active when compared to a similar population decades ago. In Germany, 42% of knee OA patients maintained involvement in sports at the time of joint replacement, and 34% continued 5 years after surgery¹¹. In addition, it has been observed that some patients started participating post-operatively in sports which they were not able to do prior surgery. For example, only a few patients were able to go hiking prior to knee joint replacement, compared to nearly 30% 5 years post-surgery¹¹. In another study performed in patients with knee OA awaiting joint replacement (mean age = 71 years), 51 and 91% of the patients considered that the KOOS sport/recreation and QOL subscales, respectively, were extremely very important⁹. Thus, the KOOS sport/recreation and QOL subscales may capture additional important information in elderly patients with knee OA. Consequently, the KOOS was evaluated in knee OA patients, including those with advanced disease progression. The KOOS was found to be valid, reliable, and responsive in OA patients before and after total joint replacement^{9,10}.

Due to the increase in large multicenter international studies and the requirement for globally meaningful epidemiologic and/or therapeutic study results, there is a need for cross-cultural adaptation and validation of health status measures. The cross-cultural adaptation of the KOOS may require not only translation but also adjustment of cultural words, idioms, and colloquialism. This process may involve substantial transformation of some items to fully capture the essence of the original concepts. Validated versions of KOOS have been currently published for use in English, Swedish⁷, Danish, German¹², Singapore English and Chinese¹³, and in numerous other languages¹⁴.

The aim of this study was to cross-culturally adapt the KOOS in French and to evaluate the psychometric properties of this adaptation, as expressed by its feasibility, construct validity, reliability, and responsiveness.

Methods

CROSS-CULTURAL ADAPTATION PROCESS

The cross-cultural adaptation was performed according to published recommendations^{15,16}. Three people (two rheumatologists and an English teacher) native in the target language independently translated the English version of the KOOS into French. A final version was written based on the consensus of the three translators. Backward translation was performed by a bilingual native English speaker (PB), blinded to the original English version. A multidisciplinary consensus committee was formed to ensure that the translation was comprehensive and verify cross-cultural equivalence of the source and final versions. The committee consisted of three rheumatologists (PO, LG, and JFM), an orthopedic surgeon (YJ), a rheumatologist and epidemiologist specialized in cross-cultural adaptation (FG), a retired rheumatologist suffering from knee OA (JS), and a native French English teacher (PG). The final version was pre-tested on 15 French patients suffering from knee OA. The patients were asked whether they fully understood all items and whether they had problems with the formulation.

EVALUATION OF THE PSYCHOMETRIC PROPERTIES OF THE FRENCH VERSION OF KOOS

Study design

Bicentric prospective study.

Patients

Two patient populations were evaluated. The medicine group was formed by all consecutive outpatients consulting for knee OA in the Rheumatology

Department of the Dijon University Hospital (France). The surgery group was recruited in the Orthopedic Surgery Department of the Marseille University Hospital (France). It was constituted by all consecutive knee OA patients waiting for total knee replacement (TKR).

The inclusion criteria were patient age of at least 40 years, and primary knee OA according to the American college of rheumatology (ACR) criteria¹⁷. Additionally, the patients in the medicine group must not have been considered as candidate for knee joint replacement while patients in the surgery group must have been presented with an indication for knee replacement. Patients had to be able to understand and complete the self-report questionnaires.

The non-inclusion criteria were the presence of other significant rheumatic disease, such as low back pain and other inferior limb joints OA, severe inflammatory arthritis as confirmed by physical examination, intra-articular use of corticosteroids within the previous 3 months and, in the medicine group, expected changes in knee OA treatment during the following 2 weeks.

Questionnaires

During the initial assessment, patients in the medicine group were asked to fill in the French version of the KOOS questionnaire and the osteoarthritis knee and hip quality of life (OAKHQOL) questionnaire. The OAKHQOL was recently validated as a specific hip and knee OA QOL instrument¹⁸. The KOOS includes five domains and 42 items. The main difference between the WOMAC and the KOOS is the inclusion of two additional domains, a five-item sport and recreation function domain, and a four-item QOL domain. In addition, four items were added to the WOMAC pain domain, and the two-item WOMAC stiffness domain was changed to a seven-item symptom domain. For each domain, scores are normalized on a 0–100 scale, 100 being the best result. The OAKHQOL contains 43 items spread over five domains (pain, physical activities, mental health, social support, and social functioning) and three independent items (sexual activity, relationships, and professional life). Scores again ranged from 0 (worst) to 100 (best).

The patients in the medicine group were given a second KOOS questionnaire which they were asked to complete 2 weeks later and to mail back, using a pre-stamped envelope.

The patients in the surgery group were asked to fill in the KOOS questionnaire pre-operatively, and during a follow-up visit, 3 months after surgery.

Missing values were handled according to KOOS and OAKHQOL guidelines. For the KOOS, when more than two of the items of a domain were missing, the score was not calculated. For OAKHQOL, when at least half of the items of a dimension were missing, the score was not calculated. In other situations, missing values were replaced by the average of values observed in the same domain for the individual.

Statistical analysis

Feasibility. Feasibility was assessed using the percentages of responses and using the floor and ceiling effects in the medicine and surgery samples. The surgery group was assessed prior to and 3 months after surgery. Floor and ceiling effects were considered to be present if more than 15% of the respondents achieved the highest or the lowest possible scores.

Construct validity. Internal consistency was assessed using Cronbach's alpha coefficient. A Cronbach's alpha equal or superior to 0.7 is generally considered as satisfactory.

Convergent and divergent construct validity was determined by comparing the results of the KOOS and OAKHQOL questionnaires. The Spearman's rank correlation was used to assess the association between domains. Coefficient correlations >0.5 , $0.5-0.35$, and <0.35 were considered as strong, moderate, and weak, respectively¹³. *A priori* hypotheses were generated for convergent (moderate to strong correlation expected) and divergent (weak correlation expected) construct validity, according to the theoretical measurement of similar or divergent construct and results of the validation studies of the KOOS questionnaires in other languages^{7,9,13}. It was hypothesized that: (1) the KOOS symptom, pain, and activity of daily life (ADL) domains would correlate strongly or moderately with the OAKHQOL pain and physical activities domains, and would correlate weakly with the other OAKHQOL domains, (2) the KOOS sports and recreation domain would correlate weakly with all OAKHQOL domains, since this domain has previously been reported as weakly correlated with all 36-item short form health survey (SF36) domains⁹ and (3) the KOOS QOL domain would correlate strongly or moderately with all OAKHQOL domains.

In addition, the baseline scores obtained in the surgery sample (prior to knee joint replacement) and the medicine sample (no indication for total joint replacement) were compared using an ANOVA (analysis of variance) (after variance homogeneity was checked). It was hypothesized that scores in the surgery sample would be statistically lower than in the medicine sample.

Reliability. The reliability of the KOOS subscales was assessed using the two questionnaires completed at a 2-week interval by the medicine sample patients. It was assumed that, using such an interval, the probability of

a spontaneous significant change in knee OA condition was low. Evaluation of the reliability used the intra-class correlation coefficient (ICC) (two way model, single measure), with 95% confidence interval (CI). An ICC of more than 0.80 is usually considered to be indicative of excellent reproducibility. In addition, the Bland and Altman representations, in which the difference between the first and the second assessments is plotted against the mean of the two assessments, were obtained¹⁹. Such representations make it possible to describe the percentage of subjects and their distribution within the 95% limits of agreements along the range of the score scale. The smallest detectable difference (SDD), which corresponds to the limits of agreement (mean change $\pm$ 1.96 standard deviation [SD] change) was obtained²⁰. The SDD indicates the smallest change that can be distinguished from the measurement error.

Responsiveness. Responsiveness was evaluated in the surgery sample by comparing the pre- and post-surgery results. The standardized response mean (SRM), i.e., the mean change between baseline and 3 months post-surgery divided by the SD of the mean change, and the effect size (ES), i.e., the mean score change between baseline and 3 months post-surgery divided by the SD of the pre-surgery values, were calculated.

The Statistical Package for the Social Sciences (SPSS) version 14.0 was used for data management and statistical analyses. Statistical significance was defined as $P < 0.05$.

Results

Although only slight differences were identified in the structure of the sentences of the 42 items between French and English KOOS versions, and between the English primary and back-translated KOOS versions, the committee discussed the questionnaire conceptual presentation, the patient comprehensibility and answer choices, before reaching a consensus.

Thirty-seven patients were included in the medicine group (mean age = 70 ± 10 years, range 45–91; 68% women). A large majority (34/37) returned their 2-week questionnaire, which allowed reliability assessment. Few individual items were missing (1.1%), and the total score could be obtained for all domains in all patients. One patient did not fill in the OAKHQOL questionnaire. In the remaining 36 patients, more than 99.5% of the items were filled correctly. Three independent items not used to calculate the different subscales of the OAKHQOL questionnaire were not fully completed. They were item 12 (ability to work, most of the included patients were retired: 88% missing data) and items 22 and 23 (related to sexual activity: 42% missing data).

Thirty patients were included in the surgery group (mean age = 71 ± 10 years, range 42–85; 73% women). Every patient filled in the KOOS questionnaire at baseline as well as at the 3-month follow-up visit. Of the individual items, 4.7% were missing. Prior to surgery, the symptoms

Table I
Internal consistency of KOOS subscales

KOOS subscales (number of items)	Cronbach's alpha coefficient
Pain (9)	0.84
Symptoms (7)	0.76
Function ADL (17)	0.93
Function sport/recreation (5)	0.84
QOL (4)	0.83

and QOL scores were obtained for all patients. The pain score was obtained for 29 patients, the ADL score for 27, and the sport and recreation score in only 25. After surgery, the pain score was obtained for all patients. The symptoms, ADL, sports and recreation, and QOL scores were not available in one, two, 10, and one patients, respectively. When reading patients' comments, it appeared that some patients did not rate the sport and recreation items because they concerned activities they had not yet engaged in daily life (running, jumping,...) at the time of post-surgery evaluation. One pain, one symptom, and three function/ADL scores were not obtained because the patients forgot to fill in one page of the 10-page questionnaire. If the forgotten pages were not taken into account, the percentage of missing individual items was 3.7%.

A floor or ceiling effect was not observed (0% of patients at the maximal and the minimal scores), except in the sport and recreation subscale. In that scale, a ceiling effect was observed in the medicine group (20.6% of patients with the worst possible score). In the surgery group, the ceiling effect was more pronounced prior to surgery (40% of patients with the worst possible score), but disappeared after joint replacement (0% with the worst possible score).

Results for internal consistency were good, with Cronbach's alpha ranging from 0.76 to 0.93 (Table I). The results of convergent and divergent construct validity are given in Table II. As expected, high or moderate correlations occurred between the KOOS and OAKHQOL scales that are intended to measure similar constructs: KOOS ADL with OAKHQOL physical activities $r_s = 0.65$ ($P < 0.001$), and KOOS pain with OAKHQOL pain $r_s = 0.42$ ($P = 0.01$). There was also an expected moderate correlation between KOOS ADL and OAKHQOL pain, and between KOOS pain and OAKHQOL physical activities. Since the OAKHQOL is a QOL questionnaire, it was not surprising to observe a good correlation between KOOS QOL and all OAKHQOL domains, except social support. The weak correlation between the KOOS sport recreation domain and all OAKHQOL domains has previously been observed with SF36

Table II
Construct validity: correlations between French KOOS and OAKHQOL subscales

KOOS OAKHQOL	Pain	Symptoms	Function ADL	Sport recreation	QOL
Physical activities	0.45 ($P = 0.008$), S or M	0.29 ($P = 0.098$), S or M	0.65 ($P < 0.001$), S or M	0.30 ($P = 0.084$), W	0.53 ($P = 0.001$), S or M
Pain	0.42 ($P = 0.014$), S or M	0.35 ($P = 0.042$), S or M	0.48 (0.004), S or M	0.20 ($P = 0.258$), W	0.54 ($P = 0.001$), S or M
Mental health	0.46 ($P = 0.007$), W	0.47 ($P = 0.005$), W	0.50 ($P = 0.002$), W	0.14 ($P = 0.44$), W	0.72 ($P < 0.001$), S or M
Social support	0.04 ($P = 0.82$), W	-0.12 ($P = 0.5$), W	0.07 ($P = 0.68$), W	0.05 ($P = 0.77$), W	0.08 ($P = 0.66$), S or M
Social functioning	0.30 ($P = 0.089$), W	0.29 ($P = 0.091$), W	0.35 ($P = 0.043$), W	0.18 ($P = 0.31$), W	0.58 ($P < 0.001$), S or M

Spearman's correlation coefficient, P -value, and expected correlation (strong, moderate or weak) are provided in each row (S: strong; M: moderate; and W: weak).

Table III

KOOS on a 0 (worst) to 100 (best) scale in 37 and 30 patients with and without indication to TKR, respectively. In the surgery group, the pain score, the ADL score, and the sport/recreation score were obtained in 29, 27, and 25 patients, respectively

KOOS subscales	Medicine group	Surgery group	P
Pain: mean (SD)	51 (16)	44 (14)	0.04
Symptoms: mean (SD)	59 (20)	49 (18)	0.02
Function ADL: mean (SD)	53 (16)	45 (14)	0.04
Function sport/recreation: mean (SD)	24 (20)	11 (16)	0.01
QOL: mean (SD)	37 (17)	23 (14)	0.001

domains⁹. In contrast, some unexpected results were obtained, in particular the high to moderate correlation between OAKHQOL mental health and all KOOS subscales except sport recreation.

The results obtained for patients in both groups are shown in Table III. As expected, the patients waiting for surgery presented with significantly lower scores in all KOOS domains.

The reproducibility of measurements of all KOOS subscales was good to excellent, with ICC ranging from 0.755 to 0.914 (Table IV). The Bland and Altman graphic representations are shown in Fig. 1. The difference between repeated measurements was not related to the mean of the two measurements. The SDD ranged from 13.4 to 21.1.

The responsiveness was high for all domains, with SRM ranging from 0.89 to 1.93 and effect size from 1.31 to 2.8 (Table V). The changes in sport and recreation domain appeared to be lower than the changes in the other subscales (Table V and Fig. 2).

Discussion

In the present work, the English version of the KOOS questionnaire was cross-culturally adapted into French. The psychometric properties of the translated version were then evaluated and found to be satisfactory. The findings will have to be confirmed by further studies evaluating other sub-populations, e.g., men and women, young and elderly, before generalizations are applicable. The French version is now available and can be downloaded on the internet¹⁴.

The psychometric properties of the French KOOS were generally similar to the original KOOS^{7,9,10}. After adaptation, the French version of KOOS seems to be a feasible instrument as illustrated by the low number of missing data. The mean scores of the QOL, and particularly the sport

and recreation function subscale were markedly lower than the scores of other KOOS subscales, as previously reported^{7,9,13}. This might support the idea that knee OA patients avoid critical situations in their daily life, such as jumping or running. However, this result might also be related to the age of the patients (mean ages = 70 and 71 years in medicine and surgery samples, respectively). Moreover, the sport and recreation domain could not be obtained in some pre- and post-joint replacement patients. The comments of some patients suggested that they did not rate some items because they corresponded to activities they never performed in daily life or, at post-surgery evaluation, because they had not yet tried to perform such activities. It is noteworthy that numerous missing data in this particular subscale had already been reported in a pre-arthroplasty population⁹. It is, however, important to include more difficult items of physical function since these are very relevant to every other patient undergoing TKR, even 5 years after surgery²¹. Except for the particular situation presented above, the percentage of missing data was very low, suggesting that the translated KOOS is suitable as a self-administered questionnaire. A formatting problem was discovered as a few patients forgot to fill in one of the pages of the questionnaire. The formatting was improved (four pages instead of 10) in order to avoid such missing data.

The internal consistency and the reliability were good and comparable to that observed in other languages^{7,12,13}. Moreover, according to the Bland and Altman graphic representation, the difference between repeated measurements was not related to the mean of the two measurements.

Comparison of the baseline data obtained in the two samples suggests that the instrument allows discrimination between patients with different levels of knee OA severity. The results of correlation support the idea that the French KOOS shows evidence of convergent and divergent construct validity. Higher correlations occurred as expected when comparing KOOS and OAKHQOL subscales measuring the same domain (pain and function). Relationships between pain and function subscales confirmed that physical function is closely related to pain in knee OA patients, in accordance with previous studies^{9,22}. The correlations between the sport and recreation function domain and all OAKHQOL domains were weak. This result might be considered surprising for OAKHQOL physical activity and pain, but was previously observed with the SF36 physical function and bodily pain domains^{9,22}. The strong to moderate correlation between the OAKHQOL mental health and all KOOS domains, except for sport and recreation, might be considered as surprising. This result was not previously observed between SF36 mental health and KOOS domains⁹. However, the OAKHQOL mental health domain

Table IV
Mean KOOS scores and reliability of French KOOS subscales

KOOS subscales	Mean KOOS score (SD)		ICC (95% CI)	SDD
	First assessment	Second assessment		
Pain	50 (15)	51 (13)	0.883 (0.78–0.94)	13.4
Symptoms	58 (19)	58 (17)	0.914 (0.834–0.956)	15.5
Function ADL	51 (14)	50 (15)	0.859 (0.736–0.927)	15.4
Function sport/recreation	21 (17)	21 (15)	0.824 (0.675–0.908)	19.6
QOL	35 (15)	34 (15)	0.755 (0.563–0.87)	21.1

Two assessments, separated by a 2-week interval, were made in 34 patients.

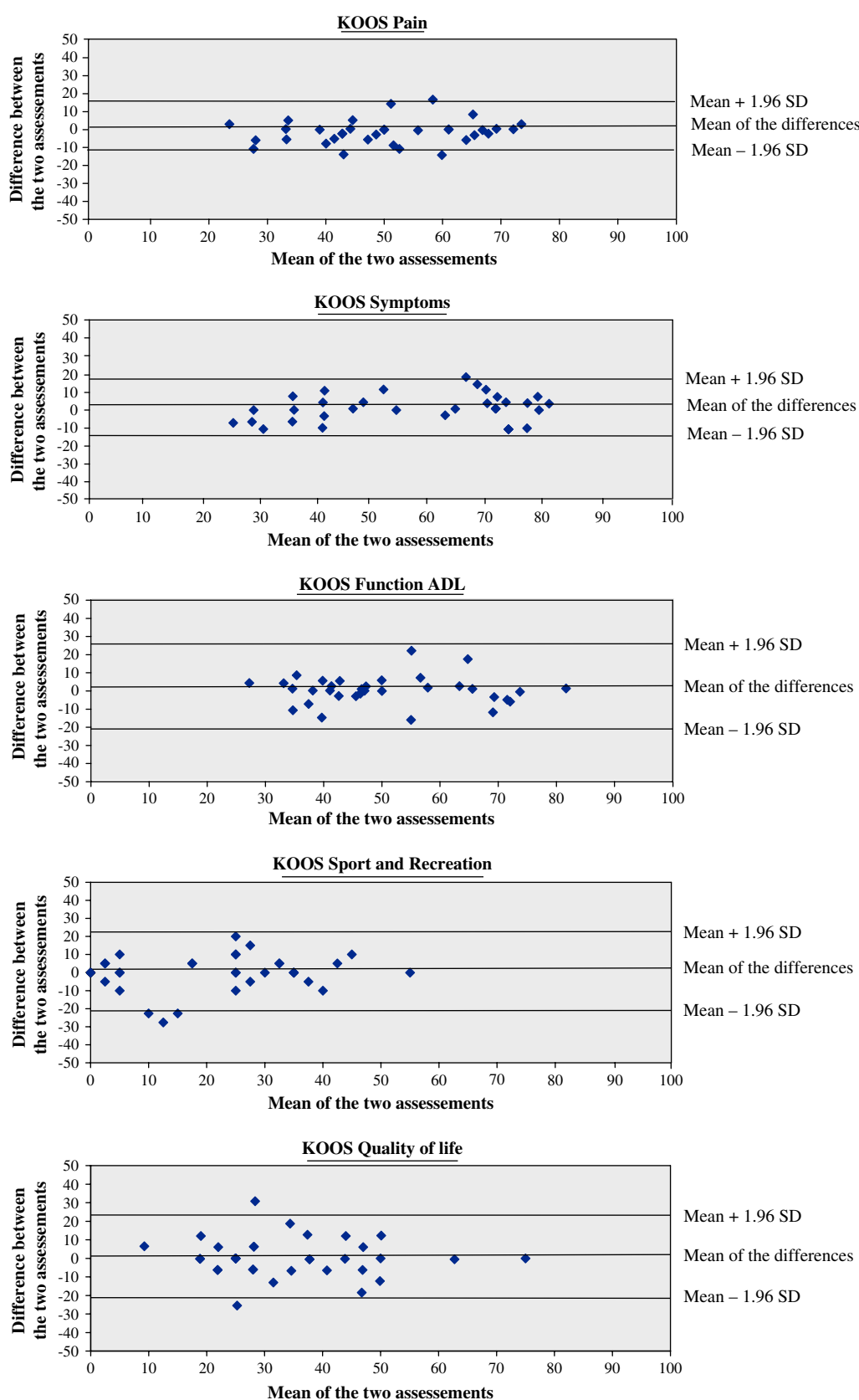


Fig. 1. Reproducibility of the French KOOS domains. Bland and Altman representations. Two assessments, separated by a 2-week interval, were made. Ninety-five percent limits of agreement correspond to the mean difference between two measurements ± 1.96 SD. (a) Pain domain, (b) symptoms domain, (c) ADL domain, (d) sports and recreation activity domain and (e) QOL domain.

Table V
Responsiveness of French KOOS subscales

KOOS subscales	ES	SRM
Pain	2.59	1.85
Symptoms	1.63	1.45
Function ADL	2.52	1.8
Function sport/recreation	1.31	0.89
QOL	2.8	1.93

Thirty patients were evaluated prior to and 3 months after TKR.

has been shown to correlate strongly with the WOMAC pain and function subscales¹⁸. Contrary to the SF36, the OAKH QOL is especially adapted to measure alterations of QOL specifically due to knee and hip OA. This might explain the differences in correlations between the KOOS domains and, on one hand the SF36, on the other hand the OAKH QOL. Interpretation of the pattern of correlations confirms that the KOOS QOL domain captures more than just pain or functional disability. Some other results were not expected but, *a posteriori*, are not so surprising. A moderate correlation was observed between the KOOS function ADL and the OAKHQOL social functioning domains. The OAKH QOL social functioning domain contains two questions which are related to the function/ADL, which can explain this result. The KOOS QOL domain was not correlated to the OAKHQOL social support domain but actually, the latter captures a dimension of QOL which is not captured by the four-item KOOS QOL domain.

The translated instrument demonstrated good responsiveness. The changes in sport and recreation domain appeared to be lower than the changes in the other subscales. The follow-up was performed only 3 months after surgery. It is unlikely that this brief 3-month interval between surgery and follow-up explains this result since a lower responsiveness of the sport and recreation, compared to other domains, has been described 6 and 12 months after surgery⁹. Moreover, in all domains, the SRM and ES observed in the present study were comparable to those described 6 and 12 months after surgery, suggesting that the choice of a 3-month period did not compromise the results. It must be pointed out, however, that total joint replacement is usually a highly effective therapy, and that the SRM and ES might not be as satisfactory if evaluated in patients treated with less efficient treatment.

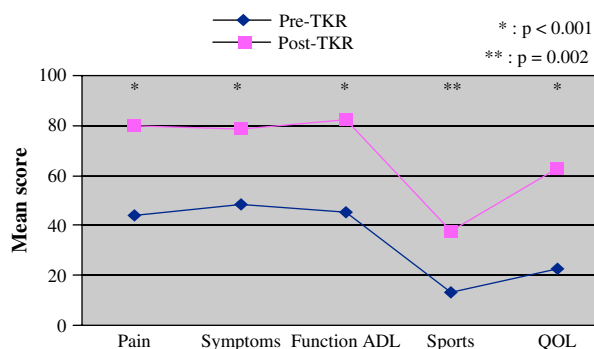


Fig. 2. KOOS domains prior to and 3 months after TKR. This scale is 0–100, worst to best.

Acknowledgments

We wish to acknowledge the other members of the multidisciplinary consensus committee: Jean Strauss, Yann Julien, and Philippe de la Grange; Philip Bastable for back translation; Christine Piroth and Hélène Chevillotte for inclusion of patients; and Djamila Zerkak for statistical help. Conflict of interest statement: The authors have no conflict of interest to declare.

References

- Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments. *Clin Rheumatol* 2005;24:29–37.
- Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. *Arch Phys Med Rehabil* 1999;80:572–9.
- Marx RG. Knee rating scales. *Arthroscopy* 2003;19:1103–8.
- Maly MR, Costigan PA, Olney SJ. Determinants of self-report outcome measures in people with knee osteoarthritis. *Arch Phys Med Rehabil* 2006;87:96–104.
- Garraff AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. *Rheumatology* (Oxford) 2004;43:1414–23.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
- Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS) – validation of a Swedish version. *Scand J Med Sci Sports* 1998;8:439–48.
- Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynon BD. Knee injury and Osteoarthritis Outcome Score (KOOS) – development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
- Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) – validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003;1:17.
- Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes* 2003;1:64.
- Huch K, Müller KAC, Stümer T, Brenner H, Puhl W, Günther KP. Sports activities 5 years after total knee or hip arthroplasty: the Ulm osteoarthritis study. *Ann Rheum Dis* 2005;64:1715–20.
- Kessler S, Lang S, Puhl W, Stove J. The Knee injury and Osteoarthritis Outcome Score – a multifunctional questionnaire to measure outcome in knee arthroplasty. *Z Orthop Ihre Grenzgeb* 2003;141:277–82.
- Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, *et al.* Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore. *Osteoarthritis Cartilage* 2006;14:1098–103.
- <http://www.koos.nu/>; 2006.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417–32.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91.
- Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al.* Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039–49.
- Rat AC, Coste J, Pouchot J, Baumann M, Spitz E, Retel-Rude N, *et al.* OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. *J Clin Epidemiol* 2005;58:47–55.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- Ravaud P, Giraudeau B, Auleley GR, Edouard-Noël R, Dougados M, Chastang C. Assessing smallest detectable change over time in continuous structural outcome measures: application to radiological change in knee osteoarthritis. *J Clin Epidemiol* 1999;12:1225–30.
- Roos EM, Toksvig-Larsen S, Nilsdotter AK. Are questions of more difficult physical function valid at 5 years after total knee replacement? *Arthritis Rheum* 2006;54(Suppl):S816.
- Paradowski PT, Englund M, Lohmander LS, Roos EM. The effect of patient characteristics on variability in pain and function over two years in early knee osteoarthritis. *Health Qual Life Outcomes* 2005;3:59.

Annexe 2. Article HOOS

Osteoarthritis and Cartilage



Cross-cultural adaptation and validation of the French version of the Hip disability and Osteoarthritis Outcome Score (HOOS) in hip osteoarthritis patients

P. Ornetti ^{†‡§*}, S. Parratte ^{||}, L. Gossec [¶], C. Tavernier ^{†‡}, J.-N. Argenson ^{||}, E.M. Roos [#],
F. Guillemin ^{††}, J.F. Maillefert ^{†‡§}

[†] Dijon University Hospital, Department of Rheumatology, Dijon F-21078, France

[‡] Bourgogne-University, Faculty of Medicine, Dijon F-21079, France

[§] INSERM U887, Dijon F-21079, France

^{||} Aix-Marseille University, Hôpital Sainte Marguerite, Department of Orthopedic Surgery, Marseille, France

[¶] Paris Descartes University, Medicine Faculty, UPRES-EA 4058, APHP, Rheumatology B Department, Cochin Hospital, Paris, France

[#] University of Southern Denmark, Institute of Sports Science and Clinical Biomechanics, Denmark

^{††} Nancy-University, Paul Verlaine Metz University, Paris Descartes University, EA4360 APEMAC, and INSERM CIC-EC CIE6, Clinical Epidemiology Department, Nancy University Hospital, France

ARTICLE INFO

Article history:

Received 18 June 2009

Accepted 11 December 2009

Keywords:

Hip osteoarthritis

HOOS

Cross-cultural adaptation

Translation

Validation

Outcomes

Pain

Disability

SUMMARY

Objective: To translate and adapt the Hip disability and Osteoarthritis Outcome Score (HOOS) into French and to evaluate the psychometric properties of this new version, by testing feasibility, internal consistency, construct validity, reliability and responsiveness, in patients with hip osteoarthritis (OA).

Methods: The French version of the HOOS was developed according to published international guidelines to ensure content validity. The new version was then evaluated in two symptomatic hip OA populations, one with no indication for joint replacement (medical group), and the other waiting for total hip replacement (THR) (surgery group). The psychometric properties assessed were *feasibility*: percentage of responses, floor and ceiling effects; *internal consistency* using Cronbach's alpha; *construct validity* by evaluating correlations with the Lequesne's index and the visual analogic scale (VAS) for pain (Pearson's correlation coefficient); *reliability*: intra-class correlation coefficient (ICC), Bland et Altman representation; *responsiveness* by comparing the results of before and 1 month after injection of hyaluronic acid (medical group) and by comparing the pre and 3 months post THR results (surgery group) by calculating standardized response mean (SRM) and effect size (ES).

Results: A total of 88 patients were recruited; 58 in the medical group (mean age = 61.8 ± 9 years, range 42–81, 70% women) and 30 in the surgery group (mean age = 67.5 ± 9 years, range 50–81, 68% women). The percentage of item responses was excellent (99%). Neither a floor nor a ceiling effect was observed, except for a ceiling effect (17.8% of patients with worst possible score) observed prior to surgery in the sport and recreation subscale. The internal consistency was good for four of the five HOOS subscales. As expected, the strongest correlations were observed between all HOOS subscales and Lequesne's index or VAS pain, indicating good construct validity. The reliability was good, with an ICC > 0.8 for all subscales. The responsiveness was good for all domains 1 month after hyaluronic acid injection (ES ranging from 0.73 to 1.86 and SRM from 0.51 to 1.04) and high for all domains 3 months after THR (ES ranging from 1.47 to 2.08 and SRM ranging from 1.97 to 3.24).

Conclusion: The French version of HOOS demonstrated good psychometric properties and appears to be useful for the evaluation of patient-relevant outcome whatever the severity of hip OA. This study provides a basis for the use of this French version of the HOOS in future clinical trials.

© 2010 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Hip osteoarthritis (OA) is one of the leading causes of chronic disability worldwide and has a significant impact on patients'

health-related quality of life (QoL). Assessment of patients with hip OA usually includes measurements of pain, functional impairment and health-related QoL using disease-specific and generic health status questionnaires^{1–4}. The most widely used OA-specific instrument is the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)⁵ that covers pain, stiffness, and functional limitation.

* Address correspondence and reprint requests to: Paul Ornetti, Department of Rheumatology, Hôpital Général, 3 rue du Faubourg Raines, 21000 Dijon, France. Tel: 33-3-80-29-37-45; Fax: 33-3-80-29-36-78.

E-mail address: paul.ornetti@chu-dijon.fr (P. Ornetti).

The Hip disability and Osteoarthritis Outcome Score (HOOS) is a simple self-administered instrument developed as an extension of the WOMAC for patients with hip OA or hip disability^{6–8}. The HOOS includes five domains, i.e., pain, other symptoms, activities of daily living (ADL), sport and recreation function (SP) and hip related QoL. The pain domain is constituted by the five questions of the original WOMAC pain domain, plus five additional questions; the other symptoms domain includes the two questions of the WOMAC stiffness domain plus three additional questions on the presence of hip noises, the difficulties to spread legs wide apart or to stride out during walk; the ADL domain corresponds to the WOMAC function (17 questions); the SP (four questions) and hip related QoL (four questions about global difficulty, lack of confidence in hip, lifestyle change or awareness of the hip problem), are newly generated domains⁶ which aim to evaluate the consequences of hip OA on more demanding activities and on QoL. A normalized score can be calculated for each domain (100 indicating no symptoms and 0 indicating extremes symptoms) and the results can be plotted in an outcome profile. The original version of the HOOS was shown to be valid, reliable, and responsive in hip OA patients and is considered useful for the evaluation of patient-relevant outcomes^{7,8}.

Due to the increase in the number of large multicenter international studies and the requirement for globally meaningful epidemiologic and/or therapeutic study results, there is a need for cross-cultural adaptation and validation of health status measures⁹. This specific process may require not only translation but also substantial transformation of some items to fully capture the essence of the original concept^{10,11}. Ideally, a final step aims to check that the translated questionnaire has kept the psychometric properties of the original one, in particular those defined by the OMER-ACT (Outcome Measures in rheumatology) filter, i.e., truth (the questionnaire reflects what it is supposed to), discrimination, which includes reliability and sensitivity to change over time, and feasibility. Validated versions of HOOS have been published for use in Danish, Dutch¹², German, English, Lithuanian and Swedish¹³. On the basis of a systematic literature search for psychometric assessment of OA-specific questionnaires, Veenhof concluded that the HOOS questionnaire was one of the top three questionnaires, with the best ratings for its psychometric properties to assess both pain and functional impairment induced by hip OA¹⁴. Thus, there is a need for translations into other languages and cross-cultural adaptations.

The aim of this study was to translate the HOOS questionnaire into French and adapt it to cultural specificities, then to evaluate the psychometric properties of the French version of the HOOS, as expressed by its feasibility, internal consistency, reliability, construct validity and responsiveness, in patients with symptomatic hip OA.

Methods

A two step procedure was used:

- Firstly, the instrument was translated and cross culturally adapted into French language.
- Secondly, the psychometric properties of the adapted instrument were studied in a prospective study by testing feasibility, internal consistency, construct validity, reliability and responsiveness.

Translation and cross-cultural adaptation process

The translation and cross-cultural adaptation from the source English version was performed according to published international recommendations^{9–11}.

Preparation

The project manager (JFM) contacted the developer (ER) to obtain permission to use and translate the instrument, to check that no other translation in French was ongoing, and to propose to participate in the work. Then, the project manager and the developer had a meeting to produce information about the conceptual basis of some items, for use by the translators.

Forward translation

Three persons (two rheumatologists, including the project manager, and one French-mother-tongue English teacher) independently translated the English version of HOOS into French. Prior to translation, the project manager provided the other translators, in particular the non-rheumatologist translator, with information about the disease, the conceptual basis of measurement, and information obtained from the developer.

Reconciliation

A final single version was obtained after a consensus meeting of the three translators.

Backward translation

The final consensus version was literally back translated into English by a bilingual native English speaker, blinded to the original English version.

Back translation review and harmonization

The project manager reviewed the back translation against the source, in order to check for discrepancies. Then, a multidisciplinary consensus committee, including the three translators, another rheumatologist (PO), an orthopedic surgeon (YJ), and a rheumatologist and epidemiologist specialized in cross-cultural adaptation (FG), was constituted. The committee had a meeting in order to perform another back translation review, as well as to conduct the harmonization step. During the meeting, the committee checked that the translation was fully comprehensive, verified cross-cultural equivalence of the source and final versions, and by consensus produced a final version.

Cognitive debriefing and review

The final version was pre-tested in 15 French outpatients visiting the rheumatology department of the Dijon University Hospital for hip OA, with characteristics comparable to those of the patients included in the second part of the study. The patients completed the questionnaire in the presence of one of the rheumatologists involved in the project. They were asked whether they fully understood all items and whether they had problems with the formulation of the items of the French version. The results of the cognitive debriefing were reviewed by the project manager. When appropriate, the initial translation was modified.

Proof reading

In the final step, the project manager checked the final translation and corrected any error. During the whole process, the participants were instructed to keep in mind that the final wording needed to be understood by non-specialists as well as individuals with a low level of education (grade 6, 10–12 years old). At any time, the project manager had the possibility to contact the developer to discuss conceptual problems.

Psychometric properties of the French version of HOOS

Data collection

Patients. Two study populations with hip OA were evaluated in a bicentric prospective study. The medical group comprised

randomly selected outpatients consulting for symptomatic hip OA in the rheumatology department of Dijon University Hospital (France). The surgery group was recruited from the orthopedic surgery department of Marseille University Hospital (France).

The inclusion criteria were age of at least 40 years, primary hip OA according to the American College of Rheumatology criteria¹⁵ and ability to understand and complete self-report questionnaires. In patients evaluated for responsiveness in the medical group, an additional inclusion criterion was the indication for intra-articular hyaluronate injection, according to the rheumatologist's usual practice. In patients evaluated for responsiveness in the surgery group, an additional inclusion criterion was the indication for total hip replacement (THR) retained by the attending orthopedic surgeon.

The exclusion criteria were the presence of other significant rheumatic diseases such as symptomatic OA in other lower limb joints, severe inflammatory arthritis and secondary hip OA and/or low back pain and the intra-articular use of corticosteroids within the previous 3 months. In patients evaluated for reliability, an additional exclusion criterion was expected changes in knee or hip OA treatment during the following 2 weeks. In patients evaluated for responsiveness in the medical group, an additional exclusion criterion was expected changes in knee or hip OA treatment during the following month, except for hyaluronate injection.

Questionnaires

Medical group. During the initial assessment, patients were asked to complete the French version of the HOOS questionnaire, Lequesne's index, and a visual analogic scale (VAS) for pain. Each HOOS item includes five answer options (Likert boxes) giving a score from 0 to 4. For each domain, scores are normalized from worst to best on a 0–100 scale. Missing values were handled according to HOOS guidelines^{6,13}: when more than two of the items of a domain were missing, the score was not calculated. In other situations, missing values were replaced by the average of values observed in the same domain for the individual. Lequesne's index^{16,17} is a disease-specific questionnaire comprising 10 items that assess pain and functional disability in hip OA. The total questionnaire, scored on a 0–24 scale, with lower scores meaning less functional impairment, is a commonly used outcome measure in lower limb OA. The VAS for pain is a widely used, valid and reliable way to measure pain intensity¹.

Patients evaluated for reliability were given a second HOOS questionnaire that was completed and returned by mail 2 weeks later, using a pre-stamped envelope. Two weeks were considered long enough to ensure that patients would not remember what they responded to the first questionnaire, and sufficiently short to ensure that no significant change in hip OA pain and disability would occur.

Patients evaluated for responsiveness were given one ultrasound-guided intra-articular injection of hyaluronic acid. The procedure was performed by the same physician (PO). The hyaluronic acid varied in nature and molecular weight since the patients presented with a specific prescription from their treating rheumatologist. The patients were given a second HOOS questionnaire, which they were asked to complete 1 month later and mail back using a pre-stamped envelope.

Surgery group. These patients were evaluated for responsiveness following THR. They were asked to complete a HOOS questionnaire before the operation, and during a follow-up visit, 3 month after surgery.

Data analysis

Feasibility. Feasibility was assessed using the percentages of responses and using the floor and ceiling effects in both the medical

and surgery group. The presence of floor and ceiling effects may influence the reliability, validity and responsiveness of an instrument. In this study, floor and ceiling effects were considered present if more than 15% of the respondents achieved the highest or lowest score.

Internal consistency. Internal consistency measures the correlations between different items on the same subscale and their degree of homogeneity. The internal consistency was assessed using Cronbach's alpha coefficient¹⁸, with 95% Feld's confidential intervals (95% CI). There are discrepancies in the literature regarding which cut-off should be used to define a satisfactory Cronbach's alpha coefficient^{19,20}. In the present study, a threshold of 0.7 was used, i.e., coefficients ≥ 0.7 were considered as satisfactory²⁰.

Construct validity. In the medical group, the construct validity was assessed by correlating the results of the five HOOS subscales with Lequesne's index and the VAS for pain. As recommended in the literature²¹, Pearson's correlation coefficients of >0.50 , 0.35 – 0.50 , and <0.35 were considered strong, moderate, and weak, respectively. *A priori* hypotheses were that all HOOS domains would strongly correlate with Lequesne's index and the pain VAS. In addition, it was hypothesised that the highest correlations would be observed between scales that are intended to measure similar constructs, i.e., VAS pain vs HOOS pain and Lequesne vs function ADL and function SP.

Reliability. The reliability of the HOOS subscales was assessed using the intra-class correlation coefficient (ICC) (two way model, single measure), with 95% CI. An ICC of more than 0.80 is usually considered an indicator of good reliability^{22,23}.

In addition, the Bland and Altman representation²⁴, in which the difference between the first and the second assessment are plotted against the mean of the two assessments, was obtained. This representation makes it possible to describe the percentage of subjects and their distribution within the 95% limits of agreements along the score scale.

The smallest detectable difference (SDD), which corresponds to the limits of agreement (mean change ± 1.96 standard deviation (SD) change) was obtained. The SDD indicates the smallest change that can be distinguished from the measurement error^{24,25}.

Responsiveness. Responsiveness was evaluated by comparing the pre and 1-month post hyaluronic acid injection results, as well as by comparing the pre and 3 months post THR results. The standardized response mean (SRM), i.e., the mean change between baseline and follow-up divided by the SD of this change, and the effect size (ES), i.e., the mean score change between baseline and follow-up divided by the SD of the baseline values, were calculated.

The Statistical Package for the Social Sciences (SPSS) version 14.0 was used for statistical analyses. The level of significance for all statistical procedures was $P < 0.05$.

Results

Translation process

Only slight differences were identified in the structure of the sentences between the three translations, as well as between the original and back-translated versions. Before reaching a consensus, the committee discussed at length the adaptation of words to maintain meaning and the correct wording to allow the questionnaire to be fully understood by all patients. A large majority of patients felt that the questionnaire was clear and easy to complete. In some instances, it was necessary to specify that the questions

Table IInternal consistency of the five HOOS subscales, $N = 88$

HOOS subscales (number of items)	Cronbach's alpha coefficient (95% CI)
Pain (10)	0.86 (0.79–0.91)
Symptoms (5)	0.66 (0.52–0.76)
Function act. daily living (17)	0.94 (0.91–0.96)
Function sport/recreation (4)	0.85 (0.78–0.90)
QoL (4)	0.82 (0.74–0.88)

95% CI: confidence interval.

were related to the previous week. Three patients thought the questionnaire contained too many questions. No modifications were made to the questionnaire following the cognitive debriefing process.

Patients' characteristics

A total of 88 patients were recruited; 58 in the medicine group (mean age = 61.8 ± 9 years, range 42–81, 70% women) and 30 in the surgery group (mean age = 67.5 ± 9 years, range 50–81, 68% women).

The samples used to evaluate the psychometric properties were as follows:

- Feasibility was obtained using the first questionnaire completed by all 88 patients.
- Internal consistency was evaluated using the first questionnaire completed by the 88 patients.
- Construct validity was evaluated on 32 patients from the medicine group from whom the VAS for pain and the Lequesne's index were obtained, in addition to the HOOS.
- Reliability was evaluated using questionnaires from 43 out of 58 patients from the medicine group. All patients completed and mailed back their second questionnaire. The remaining patients (15/58) were administrated hyaluronic injection at the first visit, so were excluded for this particular assessment.
- The evaluation of responsiveness included the 30 patients of the surgery group on one hand, 21 patients from the medicine group on the other hand. These 21 patients were the 15 who were given intra-articular hyaluronic injection at the first visit, plus six additional patients who participated to the reliability evaluation, then were treated with hyaluronic acid injection. Most patients (28 out of 30 in the surgery group and 21 out of 21 in the medicine group) completed and mailed back their second questionnaire.

Psychometric properties

Feasibility

In the medical group and the surgery group prior to THR (88 patients), few responses to HOOS items were missing (0.9% and 1.32% respectively), and the total score was obtained for all domains in all patients. No floor or ceiling effect was observed in the medical group. In the surgery sample, however, a ceiling effect was observed prior to surgery in the sport and recreation subscale (17.8% of patients with worst possible score). An additional evaluation showed that this ceiling effect had disappeared after THR (0%).

Table IIConstruct validity determined as Pearson's correlations between each subscale of the French HOOS and pain as measured by VAS and Lequesne's index, $N = 32$

HOOS	Pain	Symptoms	Function ADL	Sport recreation	QoL
Pain (VAS)	–0.67 ($P < 0.001$)	–0.49 ($P = 0.004$)	–0.71 ($P < 0.001$)	–0.58 ($P < 0.001$)	–0.55 ($P = 0.001$)
Lequesne's index	–0.68 ($P < 0.001$)	–0.54 ($P = 0.002$)	–0.69 ($P < 0.001$)	–0.82 ($P < 0.001$)	–0.51 ($P = 0.003$)

ADL: activities of daily living.

Internal consistency

Table I presents the internal consistency for the 88 patients (medical and surgery group). Cronbach's alpha ranged from 0.82 to 0.94, indicating good homogeneity, except for the symptoms subscale for which the coefficient was below 0.7 (0.66, 95% CI = 0.53–0.76). Internal consistency of the WOMAC stiffness score (which is embedded within the HOOS symptoms subscale) was poor (0.50). Cronbach's alpha was equivalent between the WOMAC pain score (0.81) and the HOOS pain subscale (0.86).

Construct validity

Table II shows the correlations between the scores of HOOS subscales, Lequesne's index and VAS pain. Strong correlations were observed between all HOOS subscales and Lequesne's index or VAS pain. As expected, the highest correlations were found between scales that measure similar constructs (Lequesne vs function ADL ($r = -0.69$) or vs function SP ($r = -0.82$)). Interestingly, the correlation between VAS pain and HOOS pain was high ($r = -0.67$), as expected, and was comparable to the correlation between VAS Pain and HOOS function ADL ($r = -0.71$).

Reliability

For all HOOS subscales, the ICCs were good, ranging from 0.83 (pain subscale) to 0.89 (sports and recreation function subscale) (Table III). The Bland and Altman graphic representations are shown in Fig. 1. The difference between repeated measurements was included within the limits of agreements in most cases, and was not related to the mean of the two measurements. The SDD ranged from 9.6 (function ADL subscale) to 16.2 (QoL subscale) (Table III). Reliability was similar between the WOMAC pain score (ICC 0.81) and the HOOS pain subscale (0.83) and between the WOMAC stiffness score (ICC 0.82) and the HOOS symptoms subscale (0.84).

Responsiveness

Surgery group. All HOOS subscale scores improved significantly ($P < 0.001$) postoperatively as compared to preoperative scores (Fig. 2a). The responsiveness was high for all domains, with SRM ranging from 1.97 to 3.24 and ES from 1.47 to 2.08 (Table IV). The QoL domain appeared to be less responsive than the other subscales (Fig. 2).

Medical group. The responsiveness was high for all domains 1 month after hyaluronic acid injection with SRM ranging from 0.51 and 1.04, and ES ranging from 0.73 to 1.86 (Table IV). The change in sport and recreation subscale appeared to be less than the change in other subscales (Fig. 2b).

The responsiveness was statistically higher ($P < 0.05$) for the HOOS pain subscale in each group (ES 1.86 for hyaluronic acid injection and ES 3.24 for THR) than for the WOMAC pain score (ES 1.25 for hyaluronic injection and 2.51 for THR). Moreover, the responsiveness was higher for the HOOS symptoms subscale than for the WOMAC stiffness score in each group, but without reaching statistical significance.

Discussion

In the present work, the English version of HOOS was cross culturally adapted into French. The psychometric properties of the

Table III

Mean HOOS scores at the first and second assessment and the test–retest reliability given as ICCs, $N = 43$

HOOS subscales	Mean HOOS score (SD) First assessment	Mean HOOS score (SD) Second assessment	ICC (95% CI)	SDD
Pain	55.3 (14.2)	52 (13.7)	0.83 (0.70–0.90)	15.1
Symptoms	54.1 (14.9)	52.9 (14.5)	0.84 (0.72–0.91)	10.5
Function ADL	53.9 (16.3)	53 (15.9)	0.86 (0.76–0.92)	9.6
Function sport/recreation	34.6 (20.4)	33.5 (20.5)	0.89 (0.80–0.94)	15.5
QoL	37.5 (19.5)	38.2 (18.3)	0.86 (0.76–0.92)	16.2

0 = worst to 100 = best.

ADL: activities of daily living.

French version were then assessed and found to be satisfactory and similar to the original HOOS. The French version of the HOOS is not copyrighted and will be downloaded for free on the KOOS website (www.KOOS.nu)¹³.

All questionnaires have been completed and the low percentage of missing data reflects the good acceptance and feasibility of the French HOOS.

The results of internal consistency were excellent for four of the five subscales and comparable to those observed in other languages^{8,12}. In line with previous validation studies, the highest Cronbach's alpha was found for the ADL subscale. Although acceptable, internal consistency for the Symptoms subscale was markedly lower (0.66) than that observed for other subscales (>0.82). This unexpected result might indicate lower inter-item

correlation in this subscale. Contrary to the subscale pain, the subscale symptoms might capture more than one concept (type of hip noises, difficulties or stiffness related to hip OA) and, thus, this multidimensional construct requires cautious interpretation of the internal consistency. However this result was not reported previously (the lowest Cronbach's alpha was found for the QoL subscale in the Swedish version of HOOS⁷ and for the Pain subscale in the Dutch version¹²). An alternative explanation might be a problem in the translation and cross-cultural adaptation into French. However, such a result was not observed in a previous study evaluating the psychometric properties of the French KOOS, in which the Symptoms subscale's internal consistency was good²⁶, while several questions of both subscales (HOOS and KOOS Symptoms) are common or very close to each other. Finally, the explanation might be variations between samples: in the present study, the 95% CI of the Symptoms subscale included 0.7.

Whatever the explanation, it is not currently possible to state that the French version of the Symptoms subscale can be considered as a homogeneous subscale. Further studies are needed to respond to this question. Waiting for these, the symptoms subscale of the French HOOS should be used at the group level rather than at the individual level (that needs a high level of internal consistency²⁷).

The mean scores of the two subscales that are not included in the WOMAC (Sport and recreation function subscale and QoL) were markedly lower than the scores of other HOOS subscales consistent with previous reports^{8,12}. This might be related to the age of our patients (mean age = 68). The test–retest reliability coefficients were high (ICC > 0.80) for all HOOS subscales, in accordance with

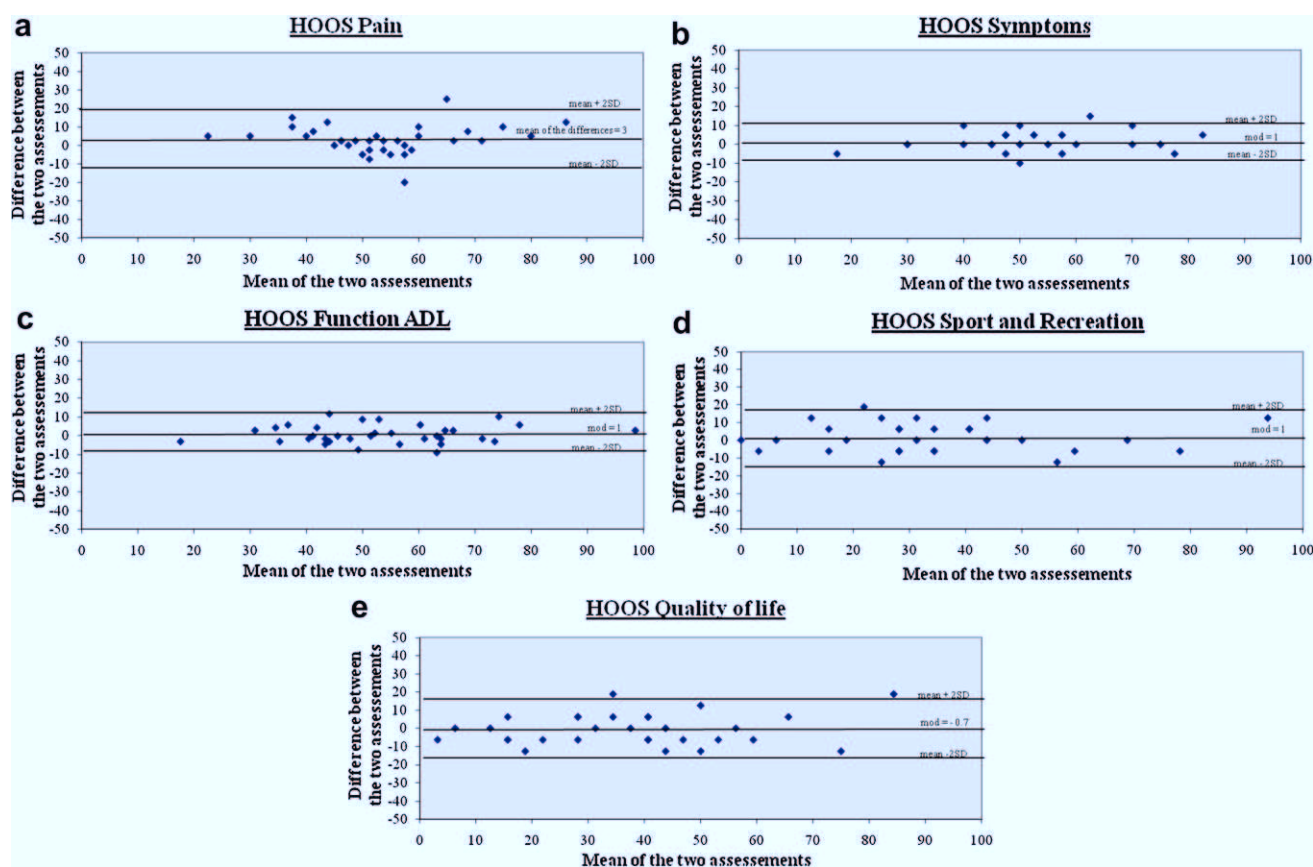


Fig. 1. Reliability of the French HOOS subscales presented as Bland and Altman representations. Two assessments, separated by a 2-week interval, were made. 95% limits of agreement correspond to the mean difference between two measurements ± 1.96 SD. (a) Pain subscale, (b) symptoms subscale, (c) activity of daily living subscale, (d) sports and recreation activity subscale and (e) QoL subscale.

Table IV
Responsiveness of French HOOS subscales

HOOS subscales	THR, N = 30				Hyaluronic acid N = 21			
	Pre mean (SD)	Post mean (SD)	ES	SRM	Pre mean (SD)	Post mean (SD)	ES	SRM
Pain	41.7 (13.8)	86.5 (16.1)	3.24	1.85	43.6 (15.7)	59.0 (17.5)	1.86	0.99
Symptoms	41.0 (17.0)	77.3 (14.2)	2.14	1.54	38.8 (16.0)	55.5 (19.7)	1.22	1.04
Function ADL	38.7 (14.7)	81.8 (13.5)	2.83	2.08	40.6 (16.2)	55.6 (18.0)	1.47	0.92
Function sport/recreation	18.3 (14.6)	64.0 (19.7)	3.11	1.90	29.5 (22.8)	41.1 (27.2)	0.73	0.51
QoL	21.4 (20.7)	61.4 (21.2)	1.97	1.47	33.9 (20.4)	48.8 (19.0)	1.00	0.73

Thirty patients were evaluated prior to and 3 months after THR.

Twenty one patients were evaluated prior to and 1 month after one ultrasound-guided intra-articular hyaluronic acid injection.

0 = worst to 100 = best.

ADL: activities of daily living.

previous studies. According to the Bland and Altman representations, the difference between repeated measurements was not related to the mean of the measurements.

Construct validity was satisfactory. All HOOS domains correlated strongly with Lequesne's index and VAS pain, with higher correlations between scales that are intended to measure similar constructs, as expected.

Responsiveness is a crucial psychometric property of a measurement instrument since high responsiveness makes it possible to reduce the number of subjects needed to demonstrate a significant difference between groups. In this study, HOOS responsiveness was good in the two evaluated samples. This result was expected in the surgery group since THR is a highly effective intervention in hip OA, and since the responsiveness of the HOOS

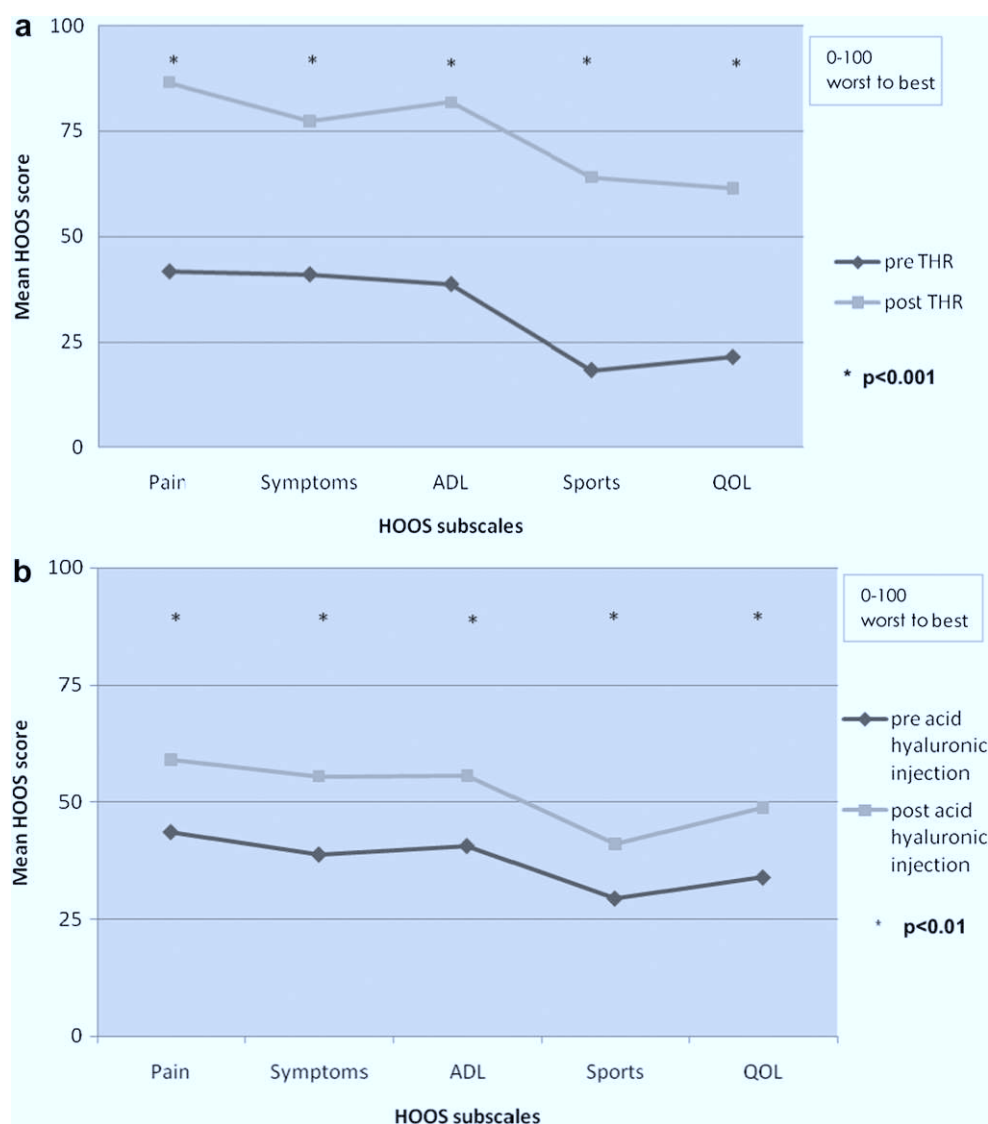


Fig. 2. Responsiveness of French HOOS. (a) HOOS subscales prior to and 3 months after THR and (b) HOOS subscales prior to and 1 month following ultrasound-guided intra-articular hyaluronic acid injection. This scale is 0–100, worst to best.

questionnaire has been found to be excellent in other THR populations in other languages^{8,12}. In contrast, to our knowledge, the responsiveness in patients treated with intra-articular hyaluronic acid injection had not been evaluated. Since this intervention is less effective than THR, and has not been found to be superior to placebo injection^{28,29}, null or low SRM and ES might have been expected, while the results demonstrated good responsiveness. However, responsiveness results of cases series studies of HA injections remains difficult to interpret due to small patient numbers, lack of controls, co-interventions, placebo response, and lack of blinding of injectors³⁰. Some of the change might have been due to a placebo effect³¹, which has been shown to be prominent in OA, particularly from injections. Moreover, in the medical group, the second assessment 1 month after HA injection might explain the high placebo response, as noticed in a previous study³² in which 25% of patients in the placebo group were considered as responders at 1 month but only 4.8% at 3 months post-treatment. Moreover, difference between groups in analgesics or NSAIDs intake might limit the ability to compare the responsiveness between THR and hyaluronic injection. Nevertheless, the aim of the present study was not to evaluate the effectiveness of hyaluronic injections in hip OA patients, but to assess the psychometric properties of a questionnaire. In that objective, the results are very satisfactory and suggest that the French HOOS can capture less prominent changes than those observed following THR and is a suitable instrument to be used in therapeutic trials. Moreover, the results of this study show that the HOOS has a higher responsiveness than the WOMAC for these patients and may be useful in evaluating hip OA and intervention outcome in different groups of patients.

This study has some limitations: firstly, participating subjects might not represent the entire spectrum of patients with hip OA as they were recruited from two University Hospitals and were likely to have more severe OA, in particular the surgery group is over representing severe OA. This might explain in part the ceiling effect detected in the surgery group prior to THR for the sport and recreation subscale, which was originally developed for younger, more active patients. This effect has been also observed in the original HOOS study⁸. Further studies evaluating other sub-populations of hip OA are needed before generalization. Secondly, the brief interval between baseline and follow-up to evaluate sensitivity to change (1 month for hyaluronic acid injection, 3 months for surgery) might appear insufficient. However, the SRM and ES observed in the present study were satisfactory and did not compromise the results.

In the last decade, the need to translate and cross culturally adapt patient reported outcome developed in English speaking countries has grown rapidly for use in international clinical trials. In the present study, a French version of the HOOS questionnaire was obtained and evaluated according to the principles of good practice for the translation and cultural adaptation process for patient reported outcomes^{9–11}. The French HOOS demonstrated good psychometric properties and appears to be useful for the evaluation of patient-relevant outcome whatever the degree of severity of hip OA. Although the internal consistency of the Symptoms domain has to be further evaluated, this study provides a basis for the use of this French version of the HOOS for future clinical trials using this instrument.

Conflict of interest

The authors P Ornetti, S Parratte, L Gossec, C Tavernier, JN Argenson, E Roos, F Guillemin have no conflict of interest. In 2007, 2008 and 2009, JF Maillefert received speaker honorarium from Abbott, Bristol Myers Squibb, Roche, and Wyeth laboratories, amounting to less than 10,000 USD.

Acknowledgments

Acknowledgments to the other members of the multidisciplinary consensus committee: Yann Julien, Philippe de la Grange, to Philip Bastable for back translation, to Christine Piroth for inclusion of patients, to Philippe d'Athis for statistical analysis.

This study was supported by Dijon University Hospital, France.

References

- Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al.* Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997 Apr;24(4):799–802.
- Boutton I, Rannou F, Jardinaud-Lopez M, Meric G, Revel M, Poiraudau S. Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year. *Osteoarthritis Cartilage* 2008 Sep;16(9):1024–31.
- Gossec L, Hawker G, Davis AM, Maillefert JF, Lohmander LS, Altman R, *et al.* OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *J Rheumatol* 2007 Jun;34(6):1432–5.
- Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. *Arch Phys Med Rehabil* 1999 May;80(5):572–9.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988 Dec;15(12):1833–40.
- Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynon BD. Knee injury and osteoarthritis outcome score (KOOS) – development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998 Aug;28(2):88–96.
- Klassbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol* 2003;32(1):46–51.
- Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003 May 30;4:10.
- Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, *et al.* Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value Health* 2005 Mar–Apr;8(2):94–104.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993 Dec;46(12):1417–32.
- Guillemin F. Cross-cultural adaptation and validation of health status measures. *Scand J Rheumatol* 1995;24(2):61–3.
- de Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra SM, Favejee M, Roos EM, *et al.* Validation of the Dutch version of the hip disability and osteoarthritis outcome score. *Osteoarthritis Cartilage* 2007 Jan;15(1):104–9.
- <http://www.koos.nu>.
- Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis

- questionnaires: a systematic review of the literature. *Arthritis Rheum* 2006 Jun 15;55(3):480–92.
15. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, *et al.* The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991 May;34(5):505–14.
 16. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation – value in comparison with other assessment tests. *Scand J Rheumatol Suppl* 1987;65:85–9.
 17. Lequesne MG. The algofunctional indices for hip and knee osteoarthritis. *J Rheumatol* 1997 Apr;24(4):779–81.
 18. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–333.
 19. George D, Mallery P. *SPSS for Windows Step by Step: a Simple Guide and Reference* 11.0 Update. 4th edn. Boston: Allyn & Bacon; 2003.
 20. Nunally JC. *Psychometric Theory*. 2nd edn. New York: McGraw-Hill; 1978.
 21. Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, *et al.* Cross-cultural adaptation and validation of Singapore English and Chinese versions of the knee injury and osteoarthritis outcome score (KOOS) in Asians with knee osteoarthritis in Singapore. *Osteoarthritis Cartilage* 2006 Nov;14(11):1098–103.
 22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159–74.
 23. Fleiss JL, Shrout P. The effects of measurement errors on some multivariate procedures. *Am J Public Health* 1977;67(12):1188–91.
 24. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 Feb 8;1(8476):307–10.
 25. Ravaud P, Giraudeau B, Auleley GR, Edouard-Noel R, Dougados M, Chastang C. Assessing smallest detectable change over time in continuous structural outcome measures: application to radiological change in knee osteoarthritis. *J Clin Epidemiol* 1999 Dec;52(12):1225–30.
 26. Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos EM, *et al.* Cross-cultural adaptation and validation of the French version of the knee injury and osteoarthritis outcome score (KOOS) in knee osteoarthritis patients. *Osteoarthritis Cartilage* 2008 Apr;16(4):423–8.
 27. McHorney C, Tarlov A. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4(4):293–307.
 28. Richette P, Ravaud P, Conrozier T, Euller-Ziegler L, Mazieres B, Maugars Y, *et al.* Effect of hyaluronic acid in symptomatic hip osteoarthritis: a multicenter, randomized, placebo-controlled trial. *Arthritis Rheum* 2009 Mar;60(3):824–30.
 29. Qvistgaard E, Christensen R, Torp-Pedersen S, Bliddal H. Intra-articular treatment of hip osteoarthritis: a randomized trial of hyaluronic acid, corticosteroid, and isotonic saline. *Osteoarthritis Cartilage* 2006 Feb;14(2):163–70.
 30. van den Bekerom MP, Lamme B, Sermon A, Mulier M. What is the evidence for viscosupplementation in the treatment of patients with hip osteoarthritis? Systematic review of the literature. *Arch Orthop Trauma Surg* 2008 Aug;128(8):815–23.
 31. Zhang W, Robertson J, Jones AC, Dieppe PA, Doherty M. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2008 Dec;67(12):1716–23.
 32. Lambert RG, Hutchings EJ, Grace MG, Jhangri GS, Conner-Spady B, Maksymowych WP. Steroid injection for osteoarthritis of the hip: a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2007 Jul;56(7):2278–87.

Annexe 3. Article KOOS-PS/HOOS-PS

Psychometric properties of the French translation of the reduced KOOS and HOOS (KOOS-PS and HOOS-PS)

P. Ornetti^{†‡§*}, A. V. Perruccio[¶], E. M. Roos^{#††}, L. S. Lohmander[#], A. M. Davis^{||‡§§|||¶¶}
and J. F. Maillefert^{†‡§}

[†] Department of Rheumatology, Dijon University Hospital, F-21079, France

[‡] Université de Bourgogne, Faculty of Medicine, Dijon, F-21078, France

[§] INSERM U887, Dijon, F-21078, France

^{||} Division of Health Care and Outcomes Research and Arthritis Community Research and Evaluation Unit, Toronto Western Research Institute, Toronto, Canada

^{¶¶} Public Health Sciences, University of Toronto, Toronto, Canada

[#] Department of Orthopaedics, Clinical Sciences Lund, Lund University, Sweden

^{††} Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Denmark

^{‡‡} Department of Physical Therapy, University of Toronto, Toronto, Canada

^{§§} Department of Rehabilitation Science, University of Toronto, Toronto, Canada

^{|||} Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada

^{¶¶} Institute of Medical Science, University of Toronto, Toronto, Canada

Summary

Objective: To evaluate the psychometric properties of the French KOOS physical function (KOOS-PS) and HOOS physical function (HOOS-PS), specifically its feasibility, reliability, construct validity, and responsiveness.

Methods: Consecutive outpatients consulting for primary knee or hip osteoarthritis (OA) in a rheumatology department were included. During the initial assessment, patients were asked to complete the Knee injury and Osteoarthritis Outcome Score (KOOS) or Hip disability and Osteoarthritis Outcome Score (HOOS) questionnaire and the Osteoarthritis Knee and Hip Quality Of Life questionnaire (OAKHQOL). The patients were given a second KOOS or HOOS questionnaire to complete and return by mail 2 weeks later.

Feasibility was assessed by calculating the percentage of missing items and the floor and ceiling effects. Test–retest reliability was evaluated using the intra-class correlation coefficient (ICC). Convergent and divergent construct validity was determined by comparing the results of the KOOS-PS or HOOS-PS and OAKHQOL questionnaires using Spearman's rank test. Responsiveness was evaluated using data obtained in other hip or knee OA patients prior to and 1 month after intra-articular hyaluronic acid injection, using standardized response mean (SRM) and effect-size (ES).

Results: Eighty-seven patients with knee OA and 50 hip OA patients were included. The KOOS-PS and HOOS-PS scores were obtained for all patients as there were no missing items. Neither a floor nor a ceiling effect was observed. The ICC of KOOS-PS and HOOS-PS was 0.861 (0.763–0.921) and 0.859 (0.725–0.929), respectively. A strong or moderate correlation was observed, as expected, between KOOS-PS, HOOS-PS, and the OAKHQOL physical activities, pain, and mental health domains. A weak correlation was observed, as expected, between KOOS-PS, HOOS-PS, and the other OAKHQOL domains, except for a moderate correlation between the KOOS-PS and social functioning. The responsiveness was demonstrated with SRM and ES of 0.80 and 0.51 (KOOS-PS), 1.10 and 0.62 (HOOS-PS), respectively.

Conclusion: The French versions of KOOS-PS and HOOS-PS are reliable, valid, and responsive questionnaires for capturing functional disability in people with knee and hip OA.

© 2009 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Key words: Knee osteoarthritis, Hip osteoarthritis, Functional evaluation, KOOS-PS, HOOS-PS, OMERACT, OARSI, Cross-cultural adaptation, Validity, Reliability, Responsiveness.

Osteoarthritis (OA) is the most common joint disease characterized by progressive destruction of cartilage, affecting to large extent weight-bearing joints, such as the knee and

hip, as well as the hand joints. The pain and disability associated with knee and hip OA have a significant impact on the patients' health-related quality of life^{1,2}. As the frequency of knee and hip OA increases as a result of life-style changes associated with higher body mass index (BMI) and less physical activity and the aging of the population, this disorder increasingly will become a major health problem. Thus, it is important to evaluate interventions that might decrease patients' disability and/or prevent or delay the progression of the disease.

*Address correspondence and reprint requests to: Dr Paul Ornetti, Department of Rheumatology, CHU Dijon Hôpital Général, 3 rue du Fb Raines, 21000 Dijon, France. Tel: 33-3-80-29-37-45; Fax: 33-3-80-29-36-78; E-mail: paul.ornetti@chu-dijon.fr
Received 23 February 2009; revision accepted 19 June 2009.

Various instruments are available to assess physical function in knee and hip OA patients^{1,3–5}. In particular, the Western Ontario and McMaster Universities Index (WOMAC) is a validated and widely used disease-specific instrument, which assesses OA-induced pain, stiffness, and functional limitations⁶. Since there were concerns that the WOMAC physical function subscale did not include items of sufficient difficulty, the Knee injury and Osteoarthritis Outcome Score (KOOS) and Hip disability and Osteoarthritis Outcome Score (HOOS) were developed as an extension of the WOMAC^{7–12}. Recently, a working group created under the auspices of OARSI (Osteoarthritis Research Society International) and OMERACT (Outcome Measures in Rheumatology Clinical Trials) considered what would be an optimal tool to evaluate physical function states that represent the progression of physical disability from early to late disease for individuals with OA of the hip and knee^{13,14}. Of primary consideration were items that represented the spectrum of disability in a short measure with appropriate measurement properties. Using the Rasch analysis and data from samples representing a spectrum of OA severity, the group developed short measures of physical function in knee and hip OA that represent the progression of physical disability, the KOOS physical function (KOOS-PS) and HOOS physical function (HOOS-PS)^{13,14}. These short measures are derived from the KOOS and HOOS and are reduced to seven (KOOS-PS) and five (HOOS-PS) items, achieving feasible, short scales with interval measurement properties that can be used as a function component of a knee and hip OA severity scoring system, covering a range of difficulty.

Due to the increase in large multicenter international studies and the requirement for globally meaningful epidemiologic and/or therapeutic study results, there is a need for cross-cultural adaptation and validation of health status measures. Moreover to assess a potential outcome measure, it is necessary to assess its psychometric properties, as defined by the OMERACT filter. The OMERACT filter¹⁵ checks that a potential outcome measure is truthful, i.e., reflects what it is supposed to reflect, and is discriminant, which includes reproducibility, and sensitivity to change, over time, and between different severity stages. The last element in the OMERACT filter refers to feasibility, which relates to time, cost, availability and is not assessed through statistics.

The aim of the present study was to evaluate the psychometric properties of the French KOOS-PS and HOOS-PS, as expressed by its feasibility, reliability, construct validity, and responsiveness.

Methods

STUDY DESIGN, PROSPECTIVE STUDY

Patients

Consecutive outpatients consulting for knee or hip OA in the rheumatology department of the Dijon University Hospital (France) were included. The inclusion criteria were patient age of at least 40 years, and primary knee or hip OA according to the American College of Rheumatology criteria¹⁶. Patients had to be able to understand and complete the self-report questionnaires. In patients evaluated for responsiveness, an additional inclusion criterion was indication for intra-articular hyaluronate injection, according to the rheumatologist's usual criterion.

The exclusion criteria were the presence of other significant rheumatic disease, such as low back pain and other lower limb joint OA, severe inflammatory arthritis as confirmed by physical examination, and intra-articular use of corticosteroids within the previous 3 months. In patients evaluated for reliability, an additional exclusion criterion was expected changes in knee or hip OA treatment during the following 2 weeks. In patients evaluated for responsiveness, an additional exclusion criterion was expected changes in knee or hip OA treatment during the following month, except for hyaluronate injection.

Questionnaires

During the initial assessment, patients were asked to complete the French versions of KOOS or HOOS questionnaires. The translation and cross-cultural adaptation process of KOOS and HOOS into French have been conducted according to recommendations and have been described elsewhere^{17,18}. Briefly, three persons (two rheumatologists and one teacher of English) native in the target language translated independently the English versions into French. A final single version was obtained after a consensus meeting. Backward translation was then performed by a bilingual native English speaker, blinded to the English original version. In the next step, a multidisciplinary consensus committee had a meeting in order to ensure that the translation was fully comprehensive and to verify cross-cultural equivalence of the source and final versions. In the last step, the final version was pre-tested among 15 French patients suffering from knee and 15 from hip OA. The KOOS-PS and HOOS-PS include seven (rising from bed, putting on socks/stockings, rising from sitting, bending to the floor, twisting/pivoting on your injured knee, kneeling, squatting) and five (descending stairs, getting in/out of bath, sitting, running, twisting/pivoting on your loaded leg) items, respectively, which were extracted in order to calculate the KOOS-PS and HOOS-PS scores. The scores were obtained as described: scored on 0–28 and 0–20 scales, respectively, then normalized on a 0–100 scale, 0 being the best^{13,14}. Patients evaluated for validity also completed the Osteoarthritis Knee and Hip Quality Of Life questionnaire (OAKHQOL) during the initial assessment. The OAKHQOL was recently validated as a specific hip and knee OA quality of life instrument¹⁹. The OAKHQOL contains 43 items spread over five domains (pain, physical activities, mental health, social support and social functioning) and three independent items (sexual activity, relationships, and professional life). Scores range from 0 (worst) to 100 (best).

Patients evaluated for reliability were given a second KOOS or HOOS questionnaire that was completed and returned by mail 2 weeks later, using a pre-stamped envelope. This length of time was chosen since it was assumed that it was sufficiently important to consider that patients would not remember what they responded to the first questionnaire, and sufficiently brief to consider that no significant change in knee or hip OA disability would occur.

Patients evaluated for responsiveness were treated with intra-articular injection of hyaluronic acid. Patients with hip OA were given one ultrasound-guided intra-articular hyaluronic acid injection. The indication for injection was based on the usual criteria of the treating rheumatologist, but the procedure was performed by the same physician (PO). The hyaluronic acid varied in nature and molecular weight since patients presented with a specific prescription from their treating rheumatologist. Patients with knee OA were given three injections at 1-week intervals. The procedure was not ultrasound-guided, and was performed by the treating rheumatologist. Again, the hyaluronic acid varied in nature and molecular weight. The patients were given a second KOOS or HOOS questionnaire which they were asked to complete 1 month after the last injection, and mail back, using a pre-stamped envelope.

For the KOOS-PS and HOOS-PS, when at least one item was missing, the score was not calculated. For OAKHQOL, when at least half of the items of a dimension were missing, the score was not calculated. When fewer items were omitted, missing values were replaced by the average of values observed in the same domain for the individual.

STATISTICAL ANALYSIS

Feasibility

Feasibility was assessed using the percentages of missing items and using the floor and ceiling effects. Floor and ceiling effects were considered present if more than 15% of the respondents achieved the highest or lowest possible scores.

Reliability

The test–retest reliability of the KOOS-PS and HOOS-PS was assessed using the two questionnaires completed at a 2-week interval. Evaluation of the reliability used the intra-class correlation coefficient (ICC) (two way model, single measure), with 95% CI (confidence interval). An ICC of more than 0.8 is usually considered to be indicative of excellent reproducibility. In addition, the Bland and Altman representation, in which the difference between the first and the second assessment is plotted against the mean of the two assessments, was obtained. Such a representation allows describing the percentage of the subjects and their distribution within the 95% limits of agreements along the range of the score scale.

Construct validity

Convergent and divergent construct validity was determined by comparing the results of the KOOS-PS or HOOS-PS and OAKHQOL questionnaires. The Spearman rank correlation was used to assess the association between domains. Coefficient correlations >0.5, 0.5–0.35, and <0.35 were considered as strong, moderate, and weak, respectively²⁰. *A priori* hypotheses were generated for convergent (moderate to strong correlation

expected) and divergent (weak correlation expected) construct validity, according to the theoretical measurement of similar or divergent constructs and to data of the literature. It was hypothesized that the KOOS-PS and HOOS-PS would correlate strongly or moderately with the OAKHQOL pain and physical activities domains, as well as with the mental health domain, since this particular domain has been shown to be strongly related to the WOMAC function subscale¹⁹. On the other hand, it was hypothesized that the KOOS-PS and the HOOS-PS would be weakly related with the other OAKHQOL domains.

Responsiveness

The responsiveness was evaluated by comparing the pre- and 1-month post-hyaluronic acid injection results. The standardized response mean (SRM), i.e., the mean change between baseline and 1 month after injection divided by the standard deviation (SD) of the mean change; and the effect-size (ES), i.e., the mean score change between baseline and 1 month after injection divided by the SD of the pre-injection values, were calculated.

The Statistical Package for the Social Sciences (SPSS) version 14.0 was used for data management and statistical analyses. Statistical significance was defined as $P < 0.05$.

Results

A total of 49 patients with knee OA (mean age = 72 ± 9 years, 71% women) were included in the reliability assessment, of which 36 were also included in the construct validity assessment. Among these 49 patients, 46 returned their 2-week questionnaires. Responsiveness was evaluated in 38 other patients, who all mailed back their 1-month post-injections questionnaires.

A total of 30 patients with hip OA (mean age = 65 ± 10 years, 74% women) were included in the reliability and construct validity assessment. All returned their 2-week questionnaires. Responsiveness was evaluated in 20 other patients, who all mailed back their 1-month post-injections questionnaires.

As no individual item was missing, the KOOS-PS and HOOS-PS scores were calculated in all patients for all assessments. Except for three independent items not used to calculate the different subscales (ability to work whereas most patients were retired, and sexual activity), the OAKHQOL questionnaires were filled in correctly, allowing calculation of the different subscales in all patients.

Neither a ceiling nor floor effect was observed as no patient had a maximal or minimal KOOS-PS or HOOS-PS score at either time of questionnaire completion.

The reproducibility of KOOS-PS and HOOS-PS was excellent (Table I). The Bland and Altman graphic representations are shown in Fig. 1. The difference between repeated measurements was included in the limits of agreements in most of cases, and was not related to the mean of the two measurements.

The results of convergent and divergent validity are shown in Table II. Eight out of 10 *a priori* hypotheses were confirmed. A strong or moderate correlation was observed, as expected, between KOOS-PS, HOOS-PS, and the OAKHQOL physical activities, pain, and mental health

domains. A weak correlation was observed, as expected, between KOOS-PS, HOOS-PS, and the other OAKHQOL domains, except for a moderate correlation between the KOOS-PS and social functioning.

The responsiveness of KOOS-PS and HOOS-PS demonstrated large effects. Following the intra-articular injection, the KOOS-PS and HOOS-PS were improved after in comparison to before intra-articular injection in a great majority of patients (33 out of 38 and 17 out of 20, respectively). The SRM and ES were 0.80 and 0.51 (KOOS-PS) and 1.1 and 0.62 (HOOS-PS), respectively (Table III).

Discussion

In the present work, the psychometric properties of the KOOS-PS and HOOS-PS were evaluated and found to be satisfactory. However, the findings will have to be confirmed by further studies evaluating other OA subpopulations before more extensive generalizability of the measures can be demonstrated. The data used in the present study were extracted from KOOS and HOOS questionnaires. Thus, the results will need confirmation with patients completing specifically the KOOS-PS and HOOS-PS questionnaires.

The results of the correlations support a convergent and divergent validity of KOOS-PS and HOOS-PS. As expected, higher correlations occurred when comparing KOOS-PS and HOOS-PS with the OAKHQOL subscale measuring the same domain (physical activity). The relationship with the OAKHQOL pain domain is in accordance with previous studies which demonstrated that functional activity and pain are related in OA patients^{9,21}. The moderate correlation between KOOS-PS and HOOS-PS and the OAKHQOL mental health domain were expected since this particular subscale has been shown to correlate with WOMAC pain, and in particular, function domains¹⁹. The relationship with OAKHQOL mental health domain was statistically significant for HOOS-PS but was not for KOOS-PS. However, the latter was close to significance ($P = 0.06$), and somewhat close coefficients of correlations were obtained, which suggest that both are correlated and that the discrepancy was due to variability in sample constitution.

A moderate correlation was observed between the KOOS-PS and the OAKHQOL social functioning domain. This might be due to the inclusion in this particular OAKHQOL domain of two questions (going out whenever one would like and have friends in whenever one would like) which might also be related to function. On the contrary, no relationship was observed between OAKHQOL social functioning domain and the HOOS-PS. Further studies are needed to assess whether this discrepancy is due to variability in sample constitution or is real.

The reliability of the reduced questionnaires was excellent especially considering the time interval between the two questionnaires. According to the Bland and Altman representations, the difference between repeated measurements was not related to the mean of the measurements. The responsiveness was good or excellent, and might even be considered as surprisingly high, given the intervention. The evaluation of responsiveness 1 month after hyaluronic acid injection might be questionable, particularly for the hip, since some studies failed to demonstrate any superiority of hyaluronic acid compared to placebo in hip OA^{22,23}. Thus, the improvement observed in the present study was probably related to a placebo effect, which has been shown to be prominent in OA, particularly from injections²⁴.

Table I
Mean KOOS-PS and HOOS-PS scores and reliability of KOOS-PS and HOOS-PS. Two assessments, separated by a 2-week interval, were made

	First assessment: mean (SD)	Second assessment: mean (SD)	ICC (95% CI)
KOOS-PS	48.3 (12.8)	49.1 (12.2)	0.861 (0.763–0.921)
HOOS-PS	51.1 (7.7)	52.1 (18)	0.859 (0.725–0.929)

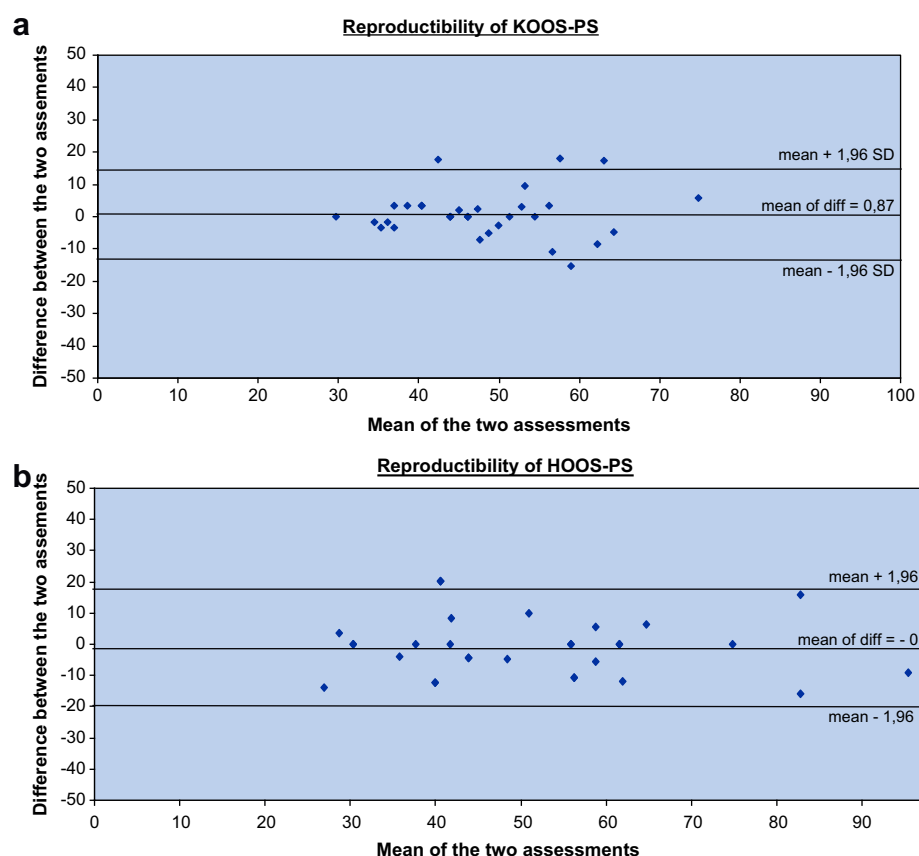


Fig. 1. Reproducibility of the French KOOS-PS and HOOS-PS: Bland and Altman representations. Two assessments, separated by a 2-week interval, were made. 95% Limits of agreement correspond to the mean difference between two measurements ± 1.96 SD. (a) KOOS-PS and (b) HOOS-PS.

However, the aim of the study was not to evaluate the efficacy of hyaluronic acid injection, but to evaluate the responsiveness of the questionnaires. Moreover, while the KOOS-PS and HOOS-PS have been shown to capture improvement induced by total joint replacement²⁵, it is of satisfaction that the questionnaire demonstrated ability to capture changes induced by a less valuable intervention. Finally, it must be pointed out that the 1-month interval was not too brief, since improvements have been described as soon as 2 weeks after injection²⁶.

Apart from reliability, validity and responsiveness, a crucial property of an instrument aimed at being used in trials is feasibility. No floor or ceiling effects were observed with the reduced questionnaires. A potential limitation of the KOOS-PS and HOOS-PS is that, in contrary to the original KOOS and HOOS in which two missing items per domain are allowed²⁷, the reduced questionnaires require that

each question is answered. Thus, another satisfying result was the absence of missing items.

A factorial analysis was not performed since the KOOS-PS and HOOS-PS were developed to ensure factorial unidimensionality^{13,14}. An additional factor analysis would not have provided further proof of this. One could object that the present data were obtained from a new sample. However, in order to ensure generalizability, the initial development of the short measures included large samples across the spectrum of disease including population-based and clinically derived samples from multiple countries.

The KOOS-PS and HOOS-PS questionnaires were developed by a task force working under the umbrella of the OARSI and OMERACT societies aiming to establish virtual criteria for total joint replacement to be used as a hard outcome in therapeutic trials²⁸. To achieve such an objective, it is critical to use instruments that represent the progression

Table II
Construct validity: correlations (Spearman's r) between KOOS-PS/
HOOS-PS and OAKHQOL subscales

OAKHQOL	KOOS-PS	HOOS-PS
Physical activities	-0.44 ($P = 0.009$)	-0.665 ($P < 0.001$)
Pain	-0.366 ($P = 0.033$)	-0.385 ($P = 0.036$)
Mental health	-0.328 ($P = 0.06$)	-0.473 ($P < 0.001$)
Social support	-0.03 ($P = 0.87$)	0.022 ($P = 0.91$)
Social functioning	-0.34 ($P = 0.05$)	-0.125 ($P = 0.5$)

Table III
Responsiveness of French KOOS-PS and HOOS-PS. Patients
were evaluated prior to and 1 month after three intra-articular
hyaluronic acid injections (KOOS-PS) or 1 month after one intra-
articular hyaluronic acid injection (HOOS-PS)

	Pre-hyaluronic acid injection: mean (SD)	Post-hyaluronic acid injection: mean (SD)	SRM	ES
KOOS-PS	48.2 (15.6)	40.3 (13.9)	0.80	0.51
HOOS-PS	51.3 (16.2)	41.3 (17.4)	1.10	0.62

of physical disability from early to late disease. The theoretical advantages of KOOS-PS and HOOS-PS are to avoid redundant items and that the Rasch model creates interval-scaled measures that provides item difficulty parameters and person functional ability parameters that are not dependant on each other. An issue might be that since the reduced questionnaires have been developed from the KOOS and HOOS, two questionnaires with numerous common items, they might not discriminate knee from hip OA. Actually, the five items of HOOS-PS are included in the KOOS questionnaire, and six out of the seven KOOS-PS items are included in the HOOS questionnaire. However, the HOOS-PS and KOOS-PS were not developed for the discrimination of the two diseases. In addition, there is only one question common to the HOOS-PS and KOOS-PS (with actually not exactly the same formulation), which means that other items of the KOOS-PS did not fit the Rasch model applied to hip OA patients and other items of the HOOS-PS did not fit the Rasch model applied to knee OA patients. This suggests that the reduced questionnaires discriminate the two affections.

A previous study of the HOOS-PS and KOOS-PS in people with total hip or knee replacement provided evidence of construct validity and responsiveness of the measures as compared to the longer WOMAC Likert 3.0 physical function subscale²⁵ but, to our knowledge, the reliability of the instruments, as well as the construct validity in patients with less advanced OA and the responsiveness following a less efficient procedure than total joint replacement, have not been assessed. Thus, apart from a validation of the French versions of the outcomes, the present work, which demonstrates a satisfying feasibility, convergent and divergent construct validity, reliability, and responsiveness can be considered as a further validation study of the new developed knee and hip OA reduced instruments aimed at assessing functional severity.

Conflict of interest

The authors have no conflict of interest.

References

- Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments. *Clin Rheumatol* 2005;24:29–37.
- Shields RK, Enloe LJ, Leo KC. Health related quality of life in patients with total hip or knee replacement. *Arch Phys Med Rehabil* 1999;80:572–9.
- Marx RG. Knee rating scales. *Arthroscopy* 2003;19:1103–8.
- Maly MR, Costigan PA, Olney SJ. Determinants of self-report outcome measures in people with knee osteoarthritis. *Arch Phys Med Rehabil* 2006;87:96–104.
- Garraff AM, Brealey S, Gillespie WJ. Patient-assessed health instruments for the knee: a structured review. *Rheumatology (Oxford)* 2004;43:1414–23.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
- Roos EM, Roos HP, Ekdahl C, Lohmander LS. Knee injury and Osteoarthritis Outcome Score (KOOS) – validation of a Swedish version. *Scand J Med Sci Sports* 1998;8:439–48.
- Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS) – development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 1998;28:88–96.
- Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) – validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes* 2003;1:17.
- Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes* 2003;1:64.
- Nilsson AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003 May 30;4:10.
- Klässbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol* 2003;32:46–51.
- Perruccio AV, Lohmander LS, Canizares M, Tennant A, Hawker GA, Conaghan PA, *et al.* The development of a short measure of physical function for knee OA. KOOS-Physical Function Short-form (KOOS-PS) – an OARSI/OMERACT initiative. *Osteoarthritis Cartilage* 2008 May;16(5):542–50.
- Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, *et al.* An OARSI/OMERACT initiative: the development of a short measure of physical function for hip OA. HOOS-Physical Function Shortform (HOOS-PS). *Osteoarthritis Cartilage* 2008 May;16(5):551–9.
- Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in Rheumatology. *J Rheumatol* 1998;25:198–9.
- Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, *et al.* Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum* 1986;29:1039–49.
- Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos E, *et al.* Cross-cultural adaptation and validation of the French version of the Knee injury and Osteoarthritis Outcome Score (KOOS) in knee osteoarthritis patients. *Osteoarthritis Cartilage* 2008;16:423–8.
- Ornetti P, Parratte S, Gossec L, Tavernier C, Argenson JN, Roos E, *et al.* Adaptation trans-culturelle et validation de la version française du HOOS (Hip Injury and Osteoarthritis Outcome Score). *Rev Rhum* 2008;75:1045 (Abstr Lu.75).
- Rat AC, Coste J, Pouchot J, Baumann M, Spitz E, Retel-Rude N, *et al.* OAKHQOL: a new instrument to measure quality of life in knee and hip osteoarthritis. *J Clin Epidemiol* 2005;58:47–55.
- Xie F, Li SC, Roos EM, Fong KY, Lo NN, Yeo SJ, *et al.* Cross-cultural adaptation and validation of Singapore English and Chinese versions of the Knee injury and Osteoarthritis Outcome Score (KOOS) in Asians with knee osteoarthritis in Singapore. *Osteoarthritis Cartilage* 2006;14:1098–103.
- Paradowski PT, Englund M, Lohmander LS, Roos EM. The effect of patient characteristics on variability in pain and function over two years in early knee osteoarthritis. *Health Qual Life Outcomes* 2005;3:59.
- Richette P, Ravaud P, Conrozier T, Euler-Ziegler L, Mazières B, Maugars Y, *et al.* Effect of hyaluronic acid in symptomatic hip osteoarthritis. *Arthritis Rheum* 2009;60:824–30.
- Fernandez Lopez JC, Ruano-Ravina A. Safety and efficacy of intra-articular hyaluronic acid in the treatment of hip osteoarthritis: a systematic review. *Osteoarthritis Cartilage* 2006;14:1306–11.
- Zhang W, Robertson J, Jones AC, Dieppe PA, Doherty M. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2008;67:1716–23.
- Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, *et al.* Comparative validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthritis Cartilage* 2009;17:843–7.
- Berg P, Olsson U. Intra-articular injection of non-animal stabilised hyaluronic acid (NASHA) for osteoarthritis of the hip: a pilot study. *Clin Exp Rheumatol* 2004;22:300–6.
- Available from: <<http://www.koos.nu>>
- Gossec L, Hawker G, Davis AM, Maillefert JF, Lohmander LS, Altman R, *et al.* OMERACT/OARSI initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. *J Rheumatol* 2007;34:1432–5.

Annexe 4. Version française du KOOS

QUESTIONNAIRE DE GENOU KOOS

DATE: _____ DATE DE NAISSANCE: _____

NOM: _____

INSTRUCTIONS

Ce questionnaire vous demande votre opinion sur votre genou. Il nous permettra de mieux connaître ce que vous ressentez et ce que vous êtes capable de faire dans votre activité de tous les jours.

Répondez à chaque question. Veuillez cocher une seule case par question. En cas de doute, cochez la case qui vous semble la plus adaptée à votre cas.

Symptômes

Ces questions concernent vos symptômes au cours des **huit derniers jours**.

S1. Est-ce que votre genou gonfle?

Jamais	Rarement	Parfois	Souvent	Tout le temps
☐	☐	☐	☐	☐

S2. Ressentez-vous des ou entendez-vous des craquements ou n'importe quel autre type de bruit en bougeant le genou?

Jamais	Rarement	Parfois	Souvent	Toujours
☐	☐	☐	☐	☐

S3. Est-ce que votre genou accroche ou se bloque en bougeant?

Jamais	Rarement	Parfois	Souvent	Toujours
☐	☐	☐	☐	☐

S4. Pouvez-vous étendre votre genou complètement?

Toujours	Souvent	Parfois	Rarement	Jamais
☐	☐	☐	☐	☐

S5. Pouvez-vous plier votre genou complètement?

Toujours	Souvent	Parfois	Rarement	Jamais
☐	☐	☐	☐	☐

Raideur

Ces questions concernent la raideur de votre genou au cours des **huit derniers jours**.

La raideur est la sensation d'avoir du mal à bouger le genou.

S6. Le matin au réveil, la raideur de votre genou est:

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

S7. Après être resté(e) assis(e), couché(e), ou au repos pendant la journée, la raideur de votre genou est:

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

Douleur

P1. Avez-vous souvent mal au genou?

Jamais

Une fois par
moisUne fois par
semaine

Tous les jours

Tout le temps

☐☐☐☐☐

Au cours des **huit derniers jours**, quelle a été l'importance de votre douleur du genou en faisant les activités suivantes?

P2. En tournant, pivotant sur votre jambe

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P3. En étendant complètement le genou

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P4. En pliant complètement le genou

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P5. En marchant sur un terrain plat

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P6. En montant ou en descendant les escaliers

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P7. Au lit la nuit

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P8. En restant assis(e) ou couché(e)

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

P9. En restant debout

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐**Fonction, vie quotidienne**

Les questions suivantes concernent ce que vous êtes capable de faire. Au cours des **huit derniers jours**, quelle a été votre difficulté pour chacune des activités suivantes?

A1. Descendre les escaliers

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

A2. Monter les escaliers

Absente

Légère

Modérée

Forte

Extrême

☐☐☐☐☐

A3. Vous relever d'une position assise

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A4. Rester debout

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A5. Vous pencher en avant pour ramasser un objet

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A6. Marcher sur un terrain plat

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A7. Monter ou descendre de voiture

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A8. Faire vos courses

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A9. Mettre vos chaussettes ou vos collants

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A10. Sortir du lit

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A11. Enlever vos chaussettes ou vos collants

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A12. Vous retourner ou garder le genou dans la même position en étant couché(e)

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A13. Entrer ou sortir d'une baignoire

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A14. Rester assis(e)

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A15. Vous asseoir ou vous relever des toilettes

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A16. Faire de gros travaux ménagers (déplacer des objets lourds, récurer les sols,...)

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

A17. Faire des petits travaux ménagers (faire la cuisine, faire la poussière,...).

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

Activités, sport et loisirs

Les questions suivantes concernent ce que vous êtes capable de faire au cours d'autres activités. Au cours des **huit derniers jours**, quelle a été votre difficulté pour les activités suivantes?

SP1. Rester accroupi(e)

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP2. Courir

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP3. Sauter

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP4. Tourner, pivoter sur votre jambe

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP5. Rester à genoux

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

Qualité de vie

Q1. Pensez-vous souvent à votre problème de genou?

Jamais	Une fois par mois	Une fois par semaine	Tous les jours	Tout le temps
☐	☐	☐	☐	☐

Q2. Avez-vous modifié votre façon de vivre pour éviter les activités qui pourraient aggraver votre problème de genou?

Pas du tout	Un peu	Modérément	Beaucoup	Totalement
☐	☐	☐	☐	☐

Q3. Est-ce qu'un manque de confiance dans votre genou vous gêne?

Pas du tout	Un peu	Modérément	Beaucoup	Totalement
☐	☐	☐	☐	☐

Q4. Finalement, êtes-vous gêné(e) par votre genou?

Pas du tout	Un peu	Modérément	Beaucoup	Extrêmement
☐	☐	☐	☐	☐

*****Merci beaucoup d'avoir répondu à ce questionnaire*****

Annexe 5. Version française du HOOS

QUESTIONNAIRE DE HANCHE HOOS

DATE: _____ DATE DE NAISSANCE: _____

NOM: _____

INSTRUCTIONS

Ce questionnaire vous demande votre opinion sur votre hanche. Il nous permettra de mieux connaître ce que vous ressentez et ce que vous êtes capable de faire dans votre activité de tous les jours.

Répondez à chaque question. Veuillez cocher une seule case par question. En cas de doute, cochez la case qui vous semble la plus adaptée à votre cas.

SYMPTOMES

Ces questions concernent vos symptômes au cours des **8 derniers jours**.

S1. Ressentez-vous des ou entendez-vous des craquements ou n'importe quel autre type de bruit provenant de votre hanche ?

Jamais	Rarement	Parfois	Souvent	Toujours
☐	☐	☐	☐	☐

S2. Quelle est votre difficulté pour écarter largement les jambes ?

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

S3. Quelle est votre difficulté pour marcher à grands pas ?

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

RAIDEUR

Ces questions concernent la raideur de votre hanche au cours des **huit derniers jours**. La raideur est la sensation d'avoir du mal à bouger la hanche.

S4. Le matin au réveil, la raideur de votre hanche est :

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

S5. Après être resté(e) assis(e), couché(e), ou au repos pendant la journée, la raideur de votre hanche est :

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

DOULEUR

D1. Avez-vous souvent mal à la hanche ?

Jamais	Une fois par mois	Une fois/semaine	Tous les jours	Tout le temps
☐	☐	☐	☐	☐

Au cours des **huit derniers jours**, quelle a été l'importance de votre douleur de hanche en faisant les activités suivantes ?

D2. En étendant complètement la hanche

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D3. En pliant complètement la hanche

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D4. En marchant sur un terrain plat

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D5. En montant ou en descendant les escaliers.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D6. Au lit la nuit

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D7. En restant assis(e) ou couché(e).

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D8. En restant debout.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D9. En marchant sur une surface dure (asphalte, béton).

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

D10. En marchant sur une surface irrégulière.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

FONCTION, VIE QUOTIDIENNE

Les questions suivantes concernent ce que vous êtes capable de faire. Au cours des **huit derniers jours**, quelle a été votre difficulté pour chacune des activités suivantes ?

F1. Descendre les escaliers.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

F2. Monter les escaliers.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

F3. Vous relever d'une position assise.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F4. Rester debout.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F5. Vous pencher en avant pour ramasser un objet

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F6. Marcher sur un terrain plat

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F7. Monter ou descendre de voiture.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F8. Faire vos courses.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F9. Mettre vos chaussettes ou vos collants

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F10. Sortir du lit

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F11. Enlever vos chaussettes ou vos collants

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F12. Vous retourner ou garder la hanche dans la même position en étant couché

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F13. Entrer ou sortir d'une baignoire.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F14. Rester assis(e).

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F15. Vous asseoir ou vous relever des toilettes.

Absente

☐

Légère

☐

Modérée

☐

Forte

☐

Extrême

☐

F16. Faire de gros travaux ménagers (déplacer des objets lourds, récurer les sols,...).

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

F17. Faire des petits travaux ménagers (faire la cuisine, faire la poussière,...).

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

ACTIVITES SPORT ET LOISIRS

Les questions suivantes concernent ce que vous êtes capable de faire au cours d'autres activités. Au cours des huit derniers jours, quelle a été votre difficulté pour les activités suivantes ?

SP1. Rester accroupi

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP2. Courir.

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP3. Tourner, pivoter sur votre jambe

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

SP4. Marcher sur une surface irrégulière

Absente	Légère	Modérée	Forte	Extrême
☐	☐	☐	☐	☐

QUALITE DE VIE

Q1. Pensez-vous souvent à votre problème de hanche ?

Jamais	Une fois par mois	Une fois/semaine	Tous les jours	Tout le temps
☐	☐	☐	☐	☐

Q2. Avez-vous modifié votre façon de vivre pour éviter les activités qui pourraient aggraver votre problème de hanche ?

Pas du tout	Un peu	Modérément	Beaucoup	Totalement
☐	☐	☐	☐	☐

Q3. Etes-ce qu'un manque de confiance dans votre hanche vous gêne ?

Pas du tout	Un peu	Modérément	Beaucoup	Totalement
☐	☐	☐	☐	☐

Q4. Finalement, êtes-vous gêné(e) par votre hanche ?

Pas du tout	Un peu	Modérément	Beaucoup	Extrêmement
☐	☐	☐	☐	☐

*****Merci beaucoup d'avoir répondu à ce questionnaire*****