



Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



Mise au point

Questionnaires pour l'ORL francophone : recensement et mise au point[☆]

S. Gargula^{1,*}, E. Babin², M.-P. Tuset³, M. Daval³, A. Mattei¹, D. Ayache³¹ Service d'otorhinolaryngologie, hôpital La Conception, CHU de Marseille, 147, boulevard Baille, 13005 Marseille, France² Service d'otorhinolaryngologie, CHU de Caen-Normandie, 3, avenue du General-Harris, 14076 Caen, France³ Service d'otorhinolaryngologie, hôpital Fondation Adolphe-de-Rothschild, 29, rue Manin, 75019 Paris, France

INFO ARTICLE

Mots clés :

Questionnaires
Patient-reported Outcome measures (PROM)
Otorhinolaryngologie

R É S U M É

Objectif. – L'utilisation d'outils d'évaluation subjectifs des patients (Patient-Reported Outcome Measures ou PROM) devient incontournable dans la pratique clinique et académique en ORL. Il est indispensable de disposer d'outils dont la traduction française a été validée. Il n'y avait actuellement aucun référencement des questionnaires disponibles en français en ORL, ni de ceux qui pourraient faire l'objet d'une adaptation transculturelle.

Méthodes. – Cette étude mandatée par le Conseil national professionnel d'ORL et la Société française d'ORL visait à recenser pour chaque surspécialité et par pathologie, les PROM répondants à l'un des critères d'inclusion : présence d'une traduction validée en français ; questionnaire non traduit mais utilisé de façon internationale (traduit dans d'autres langues et souvent cité depuis 2017), ou jugé utile subjectivement, par les experts de la surspécialité en question.

Résultats. – Un total de 103 questionnaires était répertorié. Afin de stimuler et accompagner le développement de travaux d'adaptation interculturelle et la validation statistique de questionnaires d'intérêt, notre article développe la rationalité et la méthodologie de ce type d'études.

Conclusion. – Un recensement des PROM validés en français où pouvant bénéficier d'une traduction était établi. La méthodologie de traduction et validation des PROM permettant de garantir leur fiabilité et pertinence étaient détaillées.

© 2024 Les Auteurs. Publié par Elsevier Masson SAS. Cet article est publié en Open Access sous licence CC BY (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Les questionnaires sont devenus des outils incontournables pour objectiver le ressenti du patient, sans biais subjectif de l'évaluateur, entraînant une multiplication des PROM à plusieurs échelles (macro-économique, institutionnel, pratique clinique). De façon parallèle, le nombre de questionnaires disponible en ORL a augmenté considérablement. En particulier, des questionnaires de plus en plus spécifiques ont été développés [1].

Sous l'impulsion du Conseil national professionnel d'ORL (CNPORL), un groupe de travail de la commission scientifique de la Société française d'ORL et de Chirurgie de la Face et du Cou a été créé

afin de recenser et d'apporter son soutien à l'utilisation de questionnaires de santé en pratique ORL. La première action consistait à définir les PROM et de colliger les PROM validés en langue française en ORL. Les recherches présentées dans cet article étaient réalisées de décembre 2022 à mars 2023 (et sont donc susceptibles d'avoir évolué au moment de la publication). Pour chaque surspécialité, 1 à 2 experts étaient chargés de réaliser un tri dans les questionnaires disponibles. L'identification des questionnaires était réalisée par recherche bibliographique par les experts et par consultation de l'association ou société savante de chaque surspécialité. Les critères choisis par la commission pour inclure les questionnaires étaient : présence d'une traduction validée en français ; OU questionnaire utilisé de façon internationale : traduits dans d'autres langues et cités dans PubMed depuis 2017 ; OU questionnaire ne répondant pas aux deux premiers critères, mais dont la traduction française serait néanmoins utile, à l'appréciation subjective de l'expert de chaque surspécialité. Le deuxième objectif de cet article était donner un cadre théorique sur les PROM, et des recommandations pratiques pour l'adaptation transculturelle de PROM.

DOI de l'article original : <https://doi.org/10.1016/j.anorl.2024.07.007>.

[☆] Ne pas utiliser pour citation la référence française de cet article mais celle de l'article original paru dans *European Annals of Otorhinolaryngology, Head and Neck Diseases* en utilisant le DOI ci-dessus.

* Auteur correspondant. Hôpital La Conception, 147, boulevard Baille, 13004 Marseille, France.

Adresse e-mail : stephane.gargula@gmail.com (S. Gargula).

<https://doi.org/10.1016/j.aforl.2024.05.012>

1879-7261/© 2024 Les Auteurs. Publié par Elsevier Masson SAS. Cet article est publié en Open Access sous licence CC BY (<http://creativecommons.org/licenses/by/4.0/>).

2. Résultats

2.1. Le recensement des PROM en ORL

Les tableaux de référencement des questionnaires ayant été sélectionnés dans chaque surspécialité et les références qui y sont associées sont présentés en annexe ([matériel supplémentaire](#)).

Les 103 PROM identifiés étaient divisés en 7 catégories : 28 pour otologie et audiophonologie ([Annexe 1](#)), 18 en ORL pédiatrique ([Annexe 2](#)), 22 en laryngophoniatrie ([Annexe 3](#)), 16 en cancérologie et chirurgie cervicofaciale ([Annexe 4](#)), 5 en rhinosinusologie ([Annexe 5](#)), 10 en chirurgie plastique de la face ([Annexe 6](#)), et 6 questionnaires génériques ([Annexe 7](#)).

2.2. L'utilité des PROM

A la fin des années 1980, la nécessité d'évaluer la santé de manière subjective apparaissait dans deux champs distincts : pour la santé publique et pour les essais cliniques [2]. Avec la prise en compte croissante de la dimension subjective de la santé telle qu'elle est perçue par le patient, naissait la problématique de sa mesure, et le développement des « Patient-reported outcome measures » (PROM). La première publication utilisant la qualité de vie auto-évaluée par le patient comme critère de jugement principal pour évaluer la tolérance d'un produit était une étude de 1985 [3]. Le succès de cette publication avait contribué à la création d'un groupe de travail par la Food & Drug Administration (FDA) qui a défini les PROM en 2000 [4] : « *Un PROM est un rapport sur l'état de santé d'un patient qui provient directement du patient, sans interprétation de la réponse du patient par un clinicien ou toute autre personne. Le résultat peut être mesuré en termes absolus (par exemple, la gravité d'un symptôme, d'un signe ou de l'état d'une maladie) ou en tant que changement par rapport à une mesure antérieure. Dans les essais cliniques, un instrument PROM peut être utilisé pour mesurer l'effet d'une intervention médicale sur un ou plusieurs concepts (c'est-à-dire la chose mesurée, comme un symptôme ou un groupe de symptômes, les effets sur une fonction particulière ou un groupe de fonctions, ou un groupe de symptômes ou de fonctions montrés pour mesurer la gravité d'un état de santé)* ».

Les répercussions d'une maladie ne sont pas directement proportionnelles aux symptômes, et de nombreux facteurs cognitifs et environnementaux peuvent influencer son vécu (par exemple l'impact professionnel). Deux individus avec une pathologie identique peuvent présenter des degrés de handicap très différents. L'intégration de PROM pour la mesure de qualité de vie améliore la communication entre le praticien et le patient, en cadrant la discussion sur ces problématiques [5,6]. Sur le plan académique, l'auto-évaluation des patients est devenue un des points les plus importants de l'efficacité d'une intervention médicale, et ce que la problématique initiale soit fonctionnelle, esthétique, ou carcinologique [7]. L'évaluation de l'efficacité d'une technique chirurgicale bénéficie grandement de la possibilité de mesurer la qualité de vie des patients avant et après l'intervention, et d'améliorer la comparabilité de la symptomatologie des patients [8]. En effet, la chirurgie souffre fréquemment de l'absence de comparateur direct, le double aveugle étant impossible et un essai contrôlé randomisé étant souvent difficile ou peu éthique à mettre en place (en particulier si une chirurgie factice est nécessaire, méthodologiquement) [8,9].

Suivant la même logique, les études cliniques permettant l'autorisation de mise sur le marché de nouveaux médicaments comportent de plus en plus des PROM [10,11]. Ceux-ci font désormais partie intégrante de l'évaluation de l'efficacité des traitements dans les études médicoéconomiques, comme le soutient la Haute Autorité de santé dans ses rapports méthodologiques (en particulier l'EQ-5D ; https://www.has-sante.fr/jcms/r_1499251/fr/choix-methodologiques-pour-l-evaluation-economique-a-la

[-has](#)) [12]. En parallèle des PROM, qui évaluent les résultats perçus par le patient, existent également les « Patient-Reported Experience Measures (PREM) » pour les questionnaires chargés d'évaluer « l'expérience des soins rapportée par le patient », par exemple la satisfaction vis-à-vis du parcours patient [13]. Les PREM sont davantage utilisés pour mesurer la qualité des soins au niveau des organisations de santé, plutôt qu'au niveau de la pratique clinique, et peuvent être généraux ou conçus de façon problématique-spécifique [14].

2.3. Problématiques générales et spécifiques à des populations

2.3.1. Objectifs

L'International Society for Quality of Life Research (ISOQOL) définit plusieurs types d'objectif pour les mesures de qualité de vie en santé ou « Health-related Quality of Life » (HRQOL) [15–17] : effectuer une mesure ponctuelle de la qualité de vie d'individus ; comparer la qualité de vie de différents patients ayant différentes pathologies (impact relatif des pathologies) ; mesurer l'évolution de la qualité de vie à travers le temps (essais thérapeutiques, études observationnelles) ; prédire l'évolution de la qualité de vie ; et enfin, l'utilisation en pratique clinique quotidienne, au cas par cas, avec pour objectif l'implication du patient dans sa prise en charge [18].

2.3.2. PROM généraux vs spécifiques

Les questionnaires génériques sont conçus pour mesurer toutes les dimensions de l'HRQOL, et peuvent être appliqués à tous types de populations mêmes saines [19]. Un grand avantage de ces questionnaires est qu'ils permettent de comparer différentes pathologies et leur impact sur la qualité de vie. En revanche, ils ne mesurent pas les différentes dimensions d'une pathologie spécifique, et risquent de manquer de sensibilité pour mesurer les changements d'état ou l'impact des traitements sur la qualité de vie liée à une pathologie donnée [19–21].

Les PROM spécifiques concernent les dimensions particulières d'une maladie ou d'un traitement et sont plus détaillés donc généralement plus sensibles aux modifications cliniques [20,22]. Les PROM spécifiques à une maladie sont donc plus susceptibles de faire apparaître des différences entre des thérapies concurrentes [23].

Lors de l'élaboration d'une étude scientifique, le choix du ou des questionnaires est très important. Une solution peut être d'employer dans une même étude un questionnaire spécifique et un questionnaire général, dans la limite de la pénibilité pour le patient.

2.3.3. Spécificités en pédiatrie et autres populations

L'un des principaux problèmes qui se pose en pédiatrie est la capacité limitée d'auto-évaluation de l'enfant avant un certain âge. Il est considéré qu'un enfant est capable de compléter un auto-questionnaire à partir de 8 ans [24]. Il est impératif avant cet âge de disposer d'outils d'hétéro-évaluation via un proxy (parent, professeur...) [25]. Les questionnaires d'hétéro-évaluations sont également appelés « Observer-Reported Outcomes » (ObsRO). Après cet âge, la pertinence des auto-évaluations vs hétéro-évaluations est débattue, et il semble que des discordances peuvent apparaître entre le point de vue du proxy et celui de l'enfant, en particulier pour les aspects émotionnels de la qualité de vie. Ces discordances ne sont cependant pas corrélées directement à l'âge de l'enfant, ni à son sexe [19,26].

Il peut exister par ailleurs une sur estimation de la qualité de vie de l'enfant par les parents. Ceci est mis en valeur en étudiant le « Response-shift bias », comparant l'hétéro-évaluation de la qualité de vie de manière prospective avant une intervention, et de manière rétrospective après l'intervention. L'évaluation rétrospective s'avère fréquemment pire que celle précédant l'intervention, l'amélioration clinique mettant en lumière l'état dégradé de la

qualité de vie antérieure [27]. Disposer des deux points de vue permettrait donc au final d'obtenir une vision plus globale de l'impact des pathologies sur les enfants [21], mais implique une charge plus importante pour l'évaluateur, pour les patients et pour les proxys.

L'ISPOR (The International Society for Pharmacoeconomics and Outcomes Research) Taskforce développait plusieurs recommandations pour la mesure de la qualité de vie pédiatriques [28,29] : l'utilisation des PRO conditionnée par l'âge de l'enfant ; l'utilisation de proxy lorsque nécessaire ; la confection d'outils spécifiques pour la population pédiatrique ; la confection d'outils par groupe d'âge, avec vocabulaire ou représentations imagées adaptées : une attention particulière aux problématiques d'utilisation transculturelles. Cette même Taskforce remarque qu'il est probablement impossible de fournir des seuils d'âge correspondant à toutes les situations. Elle plaçait cependant un seuil minimum pour l'auto-évaluation à 5 ans, toutes situations confondues, et plusieurs fourchettes d'âge : < 5 ans, 5–7 ans, 8–11 ans, et 12–18 ans [29]. À noter également qu'un questionnaire conçu pour étudier une large gamme d'âge est en général moins sensible aux variations de « période de vie », par exemple au cours de la transition vers l'adolescence [30]. Deux facteurs importants expliquent l'importance de ces tranches d'âge : le développement cognitif, et l'importance de placer un contexte compréhensible. En effet, des questions adaptées à une tranche d'âge peuvent avoir des performances statistiques très mauvaises sur une autre population, pour laquelle le contexte sera incompris [31]. Ces problématiques sont également des obstacles importants à l'utilisation de questionnaires conçus pour les adultes, dans une population pédiatrique.

Ce type de problématiques retrouvées particulièrement en pédiatrie peut aussi se rencontrer chez l'adulte. En particulier, les patients présentant des troubles cognitifs ou du développement peuvent présenter des difficultés à s'auto-évaluer, alors même que l'amélioration qualité de vie et de leur évaluation subjective de leurs symptômes est au centre de leur prise en charge [32]. Ces difficultés doivent impérativement être pris en compte dès la conception de l'instrument d'auto-évaluation, qui ne peut pas être intégralement substitué par une hétéro-évaluation [33]. Les questionnaires évaluant les troubles cognitifs eux-mêmes prennent généralement en compte ces aspects [34]. À l'inverse, il n'existe en général pas de questionnaire évaluant spécifiquement une pathologie complètement distincte, dans cette population (par exemple, il n'y a pas d'outil étudiant l'impact sur la qualité de vie des acouphènes chez les patients présentant une démence d'Alzheimer) [34]. La prise en charge des patients atteints de troubles mentaux nécessite également des précautions particulières. Les recommandations du National Health Service (NHS) concernant les PROM et la maladie mentale illustraient entre autres l'importance de l'accompagnement et la motivation à la réalisation des questionnaires (<https://www.rcpsych.ac.uk/improving-care/nccmh/service-design-and-development/proms-cmh-ig>).

2.4. Problématiques de recueil et d'analyse

Dans son rapport sur la place de la qualité de vie dans l'évaluation des technologies de la santé de 2018, la Haute autorité de Santé mettait l'accent sur plusieurs points axe d'amélioration des pratiques [35] :

- la qualité du recueil : les cliniciens pourraient encore être peu disposés envers les PROM (travail supplémentaire pour eux et les patients), ce qui se traduirait par des données manquantes ;
- l'analyse statistique des données : si les méthodes statistiques existent et font l'objet de nombreuses publications [36], l'analyse des données de PROM manque souvent de rigueur (p.ex. contrôle

du risque d'erreur de première espèce, gestion des données manquantes) ;

- l'interprétation des résultats : elle peut être malaisée, d'où l'enjeu de développer des outils pouvant aider à l'interprétation des scores de PROM en jugeant de la pertinence clinique des différences observées, telle que la différence minimale importante [37].

2.5. L'adaptation transculturelle et la validation de questionnaires

La première étape de tout projet incluant une évaluation par questionnaire est de chercher si une échelle adaptée aux besoins de l'étude a déjà été validée en Français. Une quantité importante d'échelles ont été traduites, mais leur version française n'a pas bénéficié d'une étude de validation proprement dite. Utiliser des échelles non validées peut avoir un impact conséquent sur la validité des résultats d'une étude. Marshall *et al.* analysaient les résultats de 300 essais randomisés portant sur l'étude des traitements médicaux en psychiatrie, et concluaient que les études utilisant des échelles non validées avaient 40 % de plus de probabilité de rapporter une efficacité du traitement par rapport aux études utilisant des échelles validées [38]. Ce chiffre était d'un tiers pour les études non médicamenteuses, et soulignait également le biais majeur que cela pouvait représenter [38].

S'il n'existe pas d'échelle validée pour la problématique, deux possibilités se présentent : créer un questionnaire de novo, ou adapter un questionnaire existant dans une autre langue. Créer un nouveau questionnaire est un processus long, rigoureux et coûteux, et pose le problème de la comparabilité avec des études étudiant des problématiques similaires mais avec des outils différents. Streiner *et al.* (Health Measurement Scales : A Practical Guide to Their Development and Use) recommandaient de ne pas créer de nouvelle échelle sauf en cas d'absolue nécessité, en soulignant que la prolifération d'échelles représente un frein à la recherche.

Dans la plupart des cas, il existe déjà une échelle, soit déjà traduite en français et non validée, soit disponible dans une autre langue, et éligible à une adaptation transculturelle. De même, modifier une échelle développée pour une problématique voisine peut être un moyen plus simple d'arriver à son but que de créer entièrement un nouveau questionnaire [39,40]. Avant de débiter l'adaptation d'un questionnaire, il est recommandé de s'assurer que celui-ci n'est pas sous Copyright, et demander sa permission (voire sa collaboration) à l'auteur de la version originale est souhaitable [39,41]. Des recommandations internationales telles que celles de l'International Test Commission (ITC guidelines) étaient créées pour homogénéiser la méthodologie des adaptations de questionnaires (et une réponse détaillée point par point à ces guidelines peut être demandée par les journaux) [42]. De même, il est recommandé d'utiliser autant que possible les guidelines COSMIN (réseau EQUATOR) pour la structuration de l'article de validation [43].

2.5.1. Le processus de traduction

L'adaptation d'un questionnaire ne se résume pas à une simple traduction [39]. En effet, une adaptation nécessite plusieurs types « d'équivalences », décrits par Herdman *et al.* [44] :

- l'équivalence conceptuelle : un même facteur étudié dans deux populations peut être interprété de manière totalement différente en fonction de la culture considérée. Par exemple, le concept de famille est très variable selon les cultures, allant de la famille nucléaire à un groupe de personnes plus étendu ;
- l'équivalence des items : les questions doivent être adaptées à la population et son mode de vie. Par exemple, une question sur le

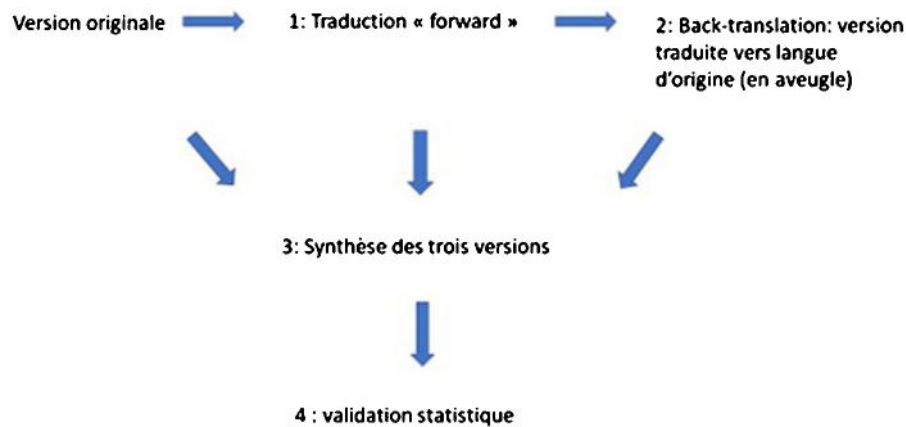


Fig. 1. Étapes successives de l'adaptation et validation transculturelle d'un questionnaire.

bruit des tondeuses à gazon dans une population urbaine n'aura pas la même portée que sur une population rurale ;

- l'équivalence sémantique : le niveau de langage doit être adapté, et les tournures de phrases compréhensibles par toute la population cible. La façon la plus simple de faire comprendre une question doit être employée quand c'est possible ;
- l'équivalence opérationnelle : le format, les instructions, le mode d'administration doivent être cohérents ;
- l'équivalence de mesure : les performances statistiques doivent être comparables entre deux versions du test, pour que les mesures elles-mêmes puissent être comparées entre elles.

La méthode de traduction la plus utilisée est la méthode du Forward-Backward [39–41]. Le principe de cette méthode est d'amplifier au maximum les erreurs et contre-sens possibles, afin de pouvoir les corriger (Fig. 1).

L'échelle est traduite et adaptée de la version originale en français par au moins deux personnes bilingues de langue maternelle française, indépendamment l'un de l'autre (ces traducteurs ne sont pas nécessairement ORL ou même du milieu médical). Ces deux versions françaises sont harmonisées par les deux traducteurs, en présence d'un investigateur, pour obtenir la version « forward ». Cette version est retraduite dans sa langue originale, de façon indépendante par deux traducteurs bilingues (si possible, la langue originale étant leur langue maternelle), en aveugle de la version originale du questionnaire, et l'un de l'autre. Ces deux versions sont harmonisées par ces deux traducteurs en présence d'un deuxième investigateur, pour obtenir la version « backward ». Les traducteurs et les deux investigateurs comparent la version « forward », « backward » et la version originale pour corriger les imprécisions linguistiques et culturelles et obtenir une première version finale française du questionnaire. À la fin de ce processus de traduction par Forward-Backward, plusieurs méthodes peuvent être utilisées pour affiner la qualité du questionnaire traduit.

Une évaluation par des experts du domaine est appréciable, et une simple notation de la qualité de la traduction de chaque item sur une échelle de 1 à 5 peut permettre d'identifier ceux pouvant possiblement être améliorés. Une évaluation de la lisibilité du questionnaire peut être faite de manière assez simple par des formules basées sur la taille des mots, le nombre de phrases, etc. [45] Celles-ci sont très faciles à trouver et utiliser directement sur internet, mais sont souvent inadaptées à l'analyse des items d'un test (on peut en revanche analyser un éventuel texte d'introduction). Enfin, une étude pilote sur un faible nombre de patients et/ou témoins est très utile pour affiner davantage le test avant l'étude sur cohorte.

2.5.2. La validation statistique d'un questionnaire

L'utilisation de questionnaires ou de traductions non validés est une source importante de biais et doit être évitée autant que possible [38]. La validation statistique est une étape essentielle pour un questionnaire de novo, pour toute nouvelle version traduite depuis une version originale, et pour toute modification conséquente (réduction du nombre d'items par exemple) [41,46]. Elle comporte nécessairement une étude prospective, sur une cohorte de patients. De nombreux paramètres peuvent et doivent être étudiés, les plus importants seront la fiabilité du test et sa validité (assimilable à sa pertinence) [46]. La fiabilité d'un test fait appel à plusieurs concepts : en premier lieu, l'homogénéité des items est indispensable pour que le test soit correctement interprétable. Celle-ci se mesure par plusieurs techniques utilisant la corrélation des items entre eux (mesure de la cohérence interne par le coefficient alpha de Cronbach par exemple), et la corrélation des items par rapport au score total (Item-total corrélation ou ITC) [47,48]. À l'homogénéité va s'adjoindre la reproductibilité des mesures. Selon ce que l'on étudie, il peut s'agir (le plus souvent) de la reproductibilité dans le temps (test-re test), ou bien de la reproductibilité interopérateurs, intra-opérateur, etc. [49]. Ces mesures font le souvent appel à des techniques statistiques de corrélation intra-classe (ICC) ou coefficients de corrélations.

Un questionnaire fiable doit cependant également être pertinent. À nouveau, plusieurs concepts méthodologiques permettent d'obtenir des informations à ce sujet, que l'on désigne parfois comme la « validité de construit ». Celle-ci signifie que le questionnaire étudie bien la problématique que l'on cherche à évaluer. La validité de convergence utilise la corrélation des résultats du test étudié avec un autre questionnaire existant qui sera considéré comme comparateur. Cependant, il n'y a en général pas d'échelle directement comparable (auquel cas en valider une nouvelle eût été redondant), et cette méthode peut être limitée. L'usage de questionnaires généraux comme comparateur peut être intéressant [50]. Un paramètre clinique (comme par exemple un test biologique, le poids, le seuil auditif.) peut être identifié comme pouvant être corrélé à la qualité de vie, et utilisé comme comparateur. Cette technique peut également être limitée par la faible corrélation retrouvée dans de nombreuses pathologies entre la qualité de vie et les paramètres objectifs. Enfin, la performance discriminatoire du test dans une population cible, par rapport à une population témoin, idéalement appariée, sera un élément méthodologique fort de la majorité des études et peut permettre de définir des seuils diagnostiques par courbe ROC.

Au total, une étude de validation de questionnaire traduit depuis une autre langue comportera une cohorte « population cible » (qui pourront recevoir deux fois le questionnaire après un intervalle

pour tester la reproductibilité), et souvent une cohorte de témoins qui servira de comparateur. Ceci permettra de produire au minimum un test de validité (performance discriminatoire), une analyse de l'homogénéité, et l'évaluation de la reproductibilité. Le nombre de cas et de témoins à inclure pourra si possible être extrapolé à partir des autres études de validations de ce questionnaire (par exemple dans d'autres langues) ou de questionnaires de longueur et complexité comparable.

2.5.3. Adaptation d'un questionnaire adulte à la population pédiatrique

Il n'y a pas de limite établie à partir de laquelle un questionnaire adulte peut être utilisé dans une population pédiatrique. Le rapport de l'ISPOR [29] donne de multiples recommandations de bonne pratique pour l'utilisation et l'adaptation pédiatrique d'un questionnaire adulte (recommandations dans les annexes, matériel supplémentaire). Il propose l'âge de 12 ans comme un âge à partir duquel l'auto-évaluation est recommandée. Cependant, cela est bien sûr variable en fonction de la complexité de l'outil, et de la sémantique utilisée [51].

3. Discussion

L'utilisation des PROM dans les publications a augmenté de façon exponentielle dans la dernière décennie en ORL, et était multipliée par 10 entre 1994 et 2013 [1]. Notre étude recensait les PROM validés les plus utilisés, ainsi que ceux jugés dignes d'intérêt par les experts de chaque spécialité. Nous présentons un cadre de la théorie et la pratique entourant l'adaptation transculturelle des questionnaires, afin de favoriser l'entreprise de telles études.

Bien que nos recherches aient mis à contribution plusieurs experts, elles peuvent ne pas être exhaustives, et des questionnaires pertinents pouvaient avoir échappé au recensement. Les auteurs rappellent que la période d'étude s'était déroulée de décembre 2022 à mars 2023, et que de nouveaux questionnaires pouvaient avoir apparu durant la période correspondant au processus de publication.

Dans un souci d'efficacité, les auteurs suggèrent aux équipes souhaitant développer des questionnaires ou s'engager dans une démarche de validation transculturelle de contacter la commission scientifique de la SFORL (s.gargula@sforl.org) afin d'éviter des travaux en parallèle qui pourraient s'avérer redondants.

4. Conclusion

L'utilisation de PROM permet d'intégrer la vision du patient et est devenue un élément crucial pour l'évaluation en santé, tant pour la prise en charge de patients en pratique clinique, que pour la réalisation d'études scientifiques. L'utilisation d'échelles validées statistiquement en français est essentielle pour limiter les biais induits par l'utilisation d'outils subjectifs. La traduction et validation des PROM comportent des étapes précises permettant de garantir leur fiabilité et pertinence.

Déclaration de liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêts.

Financement

Cette étude n'a reçu aucun financement, public ou privé.

Remerciements

Les auteurs remercient le Pr Vincent Darrouzet, Président du CNP ORL et Président de la SFORL durant la période d'étude, et le Pr Vincent Couloigner, Secrétaire Général de la SFORL pour leur initiative et pour leur soutien dans la mise en place de ce projet.

Dr Virginie Woisard (CHU de Toulouse) pour sa contribution par le recensement des PROM de laryngophoniatrie.

Mme Natacha Heutte (CHU de Caen) pour sa contribution au recensement des PROM d'oncologie-chirurgie cervicofaciale.

Pr Thomas Radulesco pour sa contribution au recensement des PROM de chirurgie plastique de la face.

Pr François Simon pour sa contribution au recensement des PROM d'ORL pédiatrique.

Pr Roger Jankowski pour sa contribution au recensement des PROM de rhinosinusologie.

Annexes 1–7. Matériel complémentaire

Le matériel complémentaire accompagnant la version en ligne de cet article est disponible sur <http://www.sciencedirect.com> et <https://doi.org/10.1016/j.aforl.2024.05.012>.

Références

- [1] Koenraads SPC, Aarts MCJ, van der Veen EL, et al. Quality of life questionnaires in otorhinolaryngology: a systematic overview. *Clin Otolaryngol* 2016;41:681–8, <http://dx.doi.org/10.1111/coa.12586>.
- [2] Akkrich M, Paterson F, Rabeharisoa V. Synthèse de la littérature sur les Patient-Reported Outcomes (2010–2019). (i3 Working Papers Series, 20-CSI-01-FR):24.
- [3] Johnson JR, Temple R. Food and Drug Administration requirements for approval of new anticancer drugs. *Cancer Treat Rep* 1985;69:1155–9.
- [4] U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft, guidance. *Health Qual Life Outcomes* 2006;4:79, <http://dx.doi.org/10.1186/1477-7525-4-79>.
- [5] Valderas JM, Kotzeva A, Espallargues M, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008;17:179–93, <http://dx.doi.org/10.1007/s11136-007-9295-0>.
- [6] Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol* 2004;22:714–24, <http://dx.doi.org/10.1200/JCO.2004.06.078>.
- [7] Nelke KH, Pawlak W, Gerber H, et al. Head and neck cancer patients' quality of life. *Adv Clin Exp Med* 2014;23:1019–27, <http://dx.doi.org/10.17219/acem/37361>.
- [8] Sokas C, Hu F, Edelen M, et al. A review of PROM implementation in surgical practice. *Ann Surg* 2022;275:85–90, <http://dx.doi.org/10.1097/SLA.0000000000005029>.
- [9] McRae C, Cherin E, Yamazaki TG, et al. Effects of perceived treatment on quality of life and medical outcomes in a double-blind placebo surgery trial. *Arch Gen Psychiatry* 2004;61:412–20, <http://dx.doi.org/10.1001/archpsyc.61.4.412>.
- [10] Gnanasakthy A, Mordin M, Evans E, et al. A review of patient-reported outcome labeling in the United States (2011–2015). *Value Health* 2017;20:420–9, <http://dx.doi.org/10.1016/j.jval.2016.10.006>.
- [11] Weis J. Assessing quality of life in clinical trials: methods and practice (2nd edition). Edited by Peter Fayers, Ron Hays. New York: Oxford University Press; 2005. p. 482 [ISBN 0-19-852769-1. *Psycho-Oncology*. 2007;16. doi:10.1002/pon.1166].
- [12] Andrade LF, Ludwig K, Goni JMR, et al. A French value set for the EQ-5D-5L. *Pharmacoeconomics* 2020;38:413–25, <http://dx.doi.org/10.1007/s40273-019-00876-4>.
- [13] Gleeson H, Calderon A, Swami V, et al. Systematic review of approaches to using patient experience data for quality improvement in healthcare settings. *BMJ Open* 2016;6:e011907, <http://dx.doi.org/10.1136/bmjopen-2016-011907>.
- [14] Reche B, McKee M, Haas M, et al. Public reporting on quality, waiting times and patient experience in 11 high-income countries. *Health Policy* 2016;120:377–83, <http://dx.doi.org/10.1016/j.healthpol.2016.02.008>.
- [15] Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–9, <http://dx.doi.org/10.7326/0003-4819-118-8-199304150-00009>.
- [16] Norman G, Cairney J. Health measurement scales: a practical guide to their development and use. *Aust N Z J Public Health* 2015;117, <http://dx.doi.org/10.1093/acprof:oso/9780199231881.003.0006>.
- [17] Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-

- centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905, <http://dx.doi.org/10.1007/s11136-012-0344-y>.
- [18] Snyder CF, Jensen RE, Segal JB, et al. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. *Med Care* 2013;51:S73–9, <http://dx.doi.org/10.1097/MLR.0b013e31829b1d84>.
- [19] Connolly MA, Johnson JA. Measuring quality of life in paediatric patients. *Pharmacoeconomics* 1999;16:605–25, <http://dx.doi.org/10.2165/00019053-199916060-00002>.
- [20] Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Health Technology Assessment (Winchester, England)* 2001;5:1–157, <http://dx.doi.org/10.3310/hta5040>.
- [21] Raat H, Mohangoo AD, Grootenhuys MA. Pediatric health-related quality of life questionnaires in clinical trials. *Curr Opin Allergy Clin Immunol* 2006;6:180–5, <http://dx.doi.org/10.1097/01.all.0000225157.67897.c2>.
- [22] Aaronson NK. Quality of life: what is it? How should it be measured? *Oncology (Williston Park)* 1988;2:69–76.
- [23] Devlin NJ, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Economics* 2010;19:886–905, <http://dx.doi.org/10.1002/hec.1608>.
- [24] Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care* 2001;39:800–12, <http://dx.doi.org/10.1097/00005650-200108000-00006>.
- [25] Varni JW, Limbers CA, Burwinkle TM. How young can children reliably and validly self-report their health-related quality of life? An analysis of 8,591 children across age subgroups with the PedsQLTM 4.0 Generic Core Scales. *Health Qual Life Outcomes* 2007;5:1, <http://dx.doi.org/10.1186/1477-7525-5-1>.
- [26] Verhey LH, Kulik DM, Ronen GM, et al. Quality of life in childhood epilepsy: what is the level of agreement between youth and their parents? *Epilepsy Behav* 2009;14:407–10, <http://dx.doi.org/10.1016/j.yebeh.2008.12.008>.
- [27] Timmerman AA, Anteunis LJ, Meesters CMG. Response-shift bias and parent-reported quality of life in children with otitis media. *Arch Otolaryngol Head Neck Surg* 2003;129:987–91, <http://dx.doi.org/10.1001/archotol.129.9.987>.
- [28] Germain N, Aballéa S, Toumi M. Measuring the health-related quality of life in young children: how far have we come? *J Mark Access Health Policy* 2019;7, <http://dx.doi.org/10.1080/20016689.2019.1618661>.
- [29] Matza LS, Patrick DL, Riley AW, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 2013;16:461–79, <http://dx.doi.org/10.1016/j.jval.2013.04.004>.
- [30] Bevans KB, Riley AW, Moon J, et al. Conceptual and methodological advances in child-reported outcomes measurement. *Expert Rev Pharmacoecon Outcomes Res* 2010;10:385–96, <http://dx.doi.org/10.1586/erp.10.52>.
- [31] Riley AW, Forrest CB, Rebok GW, et al. The Child Report Form of the CHIP-Child Edition: reliability and validity. *Med Care* 2004;42:221–31, <http://dx.doi.org/10.1097/01.mlr.0000114910.46921.73>.
- [32] Kramer JM, Schwartz A. Reducing barriers to patient-reported outcome measures for people with cognitive impairments. *Arch Phys Med Rehabil* 2017;98:1705–15, <http://dx.doi.org/10.1016/j.apmr.2017.03.011>.
- [33] American Psychological Association. Guidelines for assessment of and intervention with persons with disabilities. *Am Psychol* 2012;67:43–62, <http://dx.doi.org/10.1037/a0025892>.
- [34] Garg D, Gupta A, Agarwal A, et al. Latest trends in outcome measures in dementia and mild cognitive impairment trials. *Brain Sci* 2022;12:922, <http://dx.doi.org/10.3390/brainsci12070922>.
- [35] Évaluation des technologies de santé à la HAS : place de la qualité de vie. Haute Autorité de santé. Accessed January 23, 2023. https://www.has-sante.fr/jcms/c_2883073/fr/evaluation-des-technologies-de-sante-a-la-has-place-de-la-qualite-de-vie.
- [36] Mesbah M, Cole B, Lee ML. Statistical methods for quality of life studies: design. *Measure Anal* 2002, <http://dx.doi.org/10.1007/978-1-4757-3625-0>.
- [37] Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life—a systematic review. *J Clin Epidemiol* 2017;89:188–98, <http://dx.doi.org/10.1016/j.jclinepi.2017.06.009>.
- [38] Marshall M, Lockwood A, Bradley C, et al. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *Br J Psychiatry* 2000;176:49–252, <http://dx.doi.org/10.1192/bjp.176.3.249>.
- [39] Hall DA, Domingo SZ, Hamdache LZ, et al. A good practice guide for translating and adapting hearing-related questionnaires for different languages and cultures. *Int J Audiol* 2018;57:161–75, <http://dx.doi.org/10.1080/14992027.2017.1393565>.
- [40] Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol* 2015;68:435–41, <http://dx.doi.org/10.1016/j.jclinepi.2014.11.021>.
- [41] Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91, <http://dx.doi.org/10.1097/00007632-200012150-00014>.
- [42] Bartram D, Berberoglu G, Grégoire J, et al. ITC Guidelines for translating and adapting tests (Second Edition). *Int J Test* 2018;18:101–34, <http://dx.doi.org/10.1080/15305058.2017.1398166>.
- [43] Gagnier JJ, Lai J, Mokkink LB, et al. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res* 2021;30:2197–218, <http://dx.doi.org/10.1007/s11136-021-02822-4>.
- [44] Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res* 1998;7:323–35, <http://dx.doi.org/10.1023/a:1024985930536>.
- [45] Jindal P, MacDermid JC. Assessing reading levels of health information: uses and limitations of Flesch formula. *Educ Health (Abingdon)* 2017;30:84–8, <http://dx.doi.org/10.4103/1357-6283.210517>.
- [46] Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth* 2017;11:S80–9, <http://dx.doi.org/10.4103/sja.SJA.203.17>.
- [47] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334, <http://dx.doi.org/10.1007/BF02310555>.
- [48] Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53–5, <http://dx.doi.org/10.5116/ijme.4d4b.8d4d>.
- [49] Streiner DL, Norman GR, Cairney J. *Health measurement scales: a practical guide to their development and use*. OUP Oxford; 2014.
- [50] Bombardier C, Ware J, Russell IJ, et al. Auranofin therapy and quality of life in patients with rheumatoid arthritis. Results of a multicenter trial. *Am J Med* 1986;81:565–78, [http://dx.doi.org/10.1016/0002-9343\(86\)90539-5](http://dx.doi.org/10.1016/0002-9343(86)90539-5).
- [51] de Wit M, Delemarre-van de Waal HA, Bokma JA, et al. Monitoring and discussing health-related quality of life in adolescents with type 1 diabetes improve psychosocial well-being: a randomized controlled trial. *Diabetes Care* 2008;31:1521–6, <http://dx.doi.org/10.2337/dc08-0394>.