

# QUALITÉ DE VIE ET SANTÉ

Pr. Pascal AUQUIER  
Teodora IORDANOVA

Laboratoire de Santé Publique  
Faculté de Médecine de Marseille

Support de formation développé dans le cadre du projet INTERREG :

Coordination Euro-méditerranéenne de l'Université TETHYS.

Plate-forme numérique de Management de la Connaissance des Cultures et Territoires Méditerranéens.

Application à une approche inter-régionale de la gestion des ressources naturelles et de la santé publique dans le bassin Méditerranéen.

- 
- Commission Européenne
  - Université Euro-Méditerranéenne *TETHYS*
  - Université de la Méditerranée, Faculté de Médecine  
} (Produit par la Cellule d'Innovation Pédagogique)
  - Université Technologique de Compiègne  
} (Conçu avec le système *SCENARI*)

Avril 2004



# Table des matières

## I- Historique, Concepts et Taxonomie..... 5

Chapitre A. Introduction .....	5
1. Les raisons du recours.....	5
2. Champ d'application et Utilisation.....	8
3. Validation d'une mesure de qualité de vie.....	9
Chapitre B. Théorie de la mesure.....	10
1. Définition.....	11
2. Le paradoxe de la mesure de la qualité de vie.....	11
3. Perspective opérationnelle.....	11
Chapitre C. Les concepts mesurés.....	13
1. La santé.....	14
2. La qualité de vie liée à la santé.....	14
3. La qualité de vie.....	15
Chapitre D. La question du point de vue.....	17
1. Définition.....	17
2. Difficultés conceptuelles.....	17
3. Les différents points de vue possibles.....	18
Chapitre E. Typologies des instruments de mesure de qualité de vie.....	19
1. En fonction du champ d'application.....	19
2. En fonction de l'origine des questions.....	21
3. En fonction du type de mesure obtenue.....	22
Bibliographie.....	23

## II- Développement d'un instrument de mesure de qualité de vie..... 25

Préambule.....	25
Chapitre A. L'objectif du développement de la mesure.....	27
1. Généralités.....	27
2. Définitions.....	27

3. Discussion.....	28
Chapitre B. Le contenu.....	29
1. Analyse de contenu.....	29
2. Analyse de la littérature.....	29
3. Jugement d'experts.....	29
Chapitre C. Construction du questionnaire.....	30
1. Construction des questions.....	30
2. Adaptation - Ecriture.....	35
3. Revue des énoncés.....	35
Chapitre D. Tri des questions.....	37
1. Nombre de sujets.....	37
2. Réalisation.....	37
3. Réalisation de l'enquête et analyse.....	38
Chapitre E. Validation, Acceptabilité et recommandation.....	39
1. Validation.....	39
2. Acceptabilité.....	42
3. Ecriture des recommandations d'utilisation pratique.....	42
Bibliographie.....	42

### **III- Adaptation trans-culturelle d'instruments existants..... 44**

Préambule.....	44
Chapitre A. Concept d'équivalence.....	45
1. Equivalence conceptuelle.....	45
2. Equivalence opérationnelle.....	46
3. Equivalence de construit ou d'échelle.....	46
4. Equivalence métrique.....	46
Chapitre B. Les différentes approches.....	46
1. Approche ethnocentrique.....	47
2. Approche pragmatique.....	47
3. Approche emic plus etic.....	47
4. Approche analogique.....	47
Chapitre C. Déroulement du processus de traduction.....	48
1. Traduction et retro-traduction.....	48
2. Etablissement du questionnaire final.....	49
3. Pré-test.....	49
4. Harmonisation internationale.....	50
Chapitre D. Bibliographie.....	50
1. Références, citées dans le texte : lecture recommandée.....	50

2. Articles cités dans le texte .....	51
3. Non cités dans le texte : pour plus d'information. ....	52
4. Schéma : processus de traduction .....	53

# Historique, Concepts et Taxonomie

## Chapitre A. Introduction

### 1. Les raisons du recours

Les recherches sur la mesure de la qualité de vie se développent rapidement. Le nombre des études et des articles qui lui sont consacrés va croissant. La véritable nouveauté ne réside pas tant dans l'intérêt que portent les professionnels de la santé à la qualité de vie de leurs patients, que dans leur volonté d'user d'instruments de mesure qui permettent de quantifier cette qualité de vie en fonction des réponses des patients (1). L'intérêt actuel que le monde médical (médecins et patients) porte à la mesure de la qualité de vie s'explique de plusieurs façons.

#### *Les progrès thérapeutiques*

Tout d'abord, les progrès thérapeutiques, en particulier dans le domaine des maladies infectieuses, ont participé à l'accroissement de la longévité observé dans les pays industrialisés et à la recrudescence corrélative des maladies de nature dégénérative, dont l'étiologie est mal comprise, et contre lesquelles la médecine ne dispose pas toujours d'un traitement curatif. Dans ces circonstances, les traitements disponibles sont jugés à partir de leur capacité à enrayer les conséquences, non seulement physiques, mais aussi mentales et sociales des maladies chroniques, autrement dit, de leur capacité à restaurer ou à préserver la qualité de vie.

En outre, certains médicaments qui permettent de lutter contre une affection chronique banale et peu symptomatique telle que l'hypertension artérielle, peuvent avoir des effets indésirables (par exemple sur la sexualité). Nul ne contestera qu'il est raisonnable de s'interroger sur l'impact, en termes de qualité de vie, des traitements non curatifs des maladies chroniques. De plus, certaines classes de traitements dont les effets, évalués selon des critères traditionnels, sont controversés, par exemple les traitements à base d'hormone de croissance, peuvent trouver une certaine légitimité s'il est prouvé qu'ils ont un effet bénéfique sur la qualité de vie des sujets auxquels ils sont administrés (2). La multiplication des spécialités pharmaceutiques concurrentes pour traiter une même situation pathologique a, quant à elle, suscité la recherche d'instruments capables de mettre en évidence des différences qui n'étaient pas signalées par les indicateurs traditionnels dont disposaient les cliniciens (symptômes, guérison, taux de mortalité, etc.). Quand l'efficacité clinique des traitements est similaire, le prescripteur prend davantage en compte la sévérité de leurs effets secondaires et de leurs conséquences sur la vie quotidienne des patients. C'est à ce titre que certains publicitaires ont mis l'accent sur le caractère non préjudiciable de tel ou tel médicament en termes de qualité de vie.

### *Les limites du concept de " progrès médical continu "*

La multiplication des interventions invasives et des traitements présentant de graves effets secondaires, par exemple en cancérologie, ont conduit plusieurs chercheurs à s'interroger sur l'efficacité d'un certain nombre d'actions médicales (3). L'accélération de l'innovation médicale, par exemple de l'imagerie, et des investissements qu'elle suscite, sans que le bénéfice pour les patients soit toujours apparent, ont conduit une partie du public à mettre en doute la pertinence de l'excellence technique à tout prix. Le mythe d'un progrès médical continu et nécessairement bon a été battu en brèche.

### *Le mouvement consumériste*

Les patients se sont, de plus en plus, conduits en consommateurs et ont demandé à être associés aux décisions les concernant. La multiplication des procès contre les médecins, surtout aux USA, témoigne de cette évolution. Elle témoigne aussi des insuffisances de la communication entre les médecins et leurs malades en ce qui concerne la formulation des objectifs des soins médicaux, et l'évaluation des réalisations de la médecine. A cet égard, les mesures "objectives" de la qualité de vie peuvent participer à la défense de la profession médicale contre ces attaques.

### *La maîtrise économique du système de santé*

Quant aux instances chargées du financement, elles ont cherché un moyen de s'assurer du bien-fondé de leurs dépenses. L'explosion, apparemment incontrôlable, des budgets médicaux et la nécessité pressante de maîtriser les coûts, tout en maintenant la qualité des soins, expliquent suffisamment l'intérêt pour les mesures de qualité de vie liée à la santé.

### *L'évolution du concept de santé*

Enfin, la définition même des objectifs des actions médicales a évolué. En témoigne la fameuse définition donnée par l'Organisation Mondiale de la Santé (OMS) (4), où la santé est pensée positivement comme "un état de complet bien-être physique, psychologique et social, et pas seulement comme l'absence de maladie". Aussi la médecine ne peut-elle plus négliger de considérer la qualité de vie de ses patients.

### *Les limites du jugement médical " objectif "*

L'intérêt que cliniciens et chercheurs en santé publique portent aux mesures de qualité de vie liée à la santé s'explique en partie par la prise de conscience qu'il est nécessaire de tenir compte des perceptions et des préférences des patients, en matière de décisions de santé. Même si ce sont les médecins qui prescrivent les soins, ce sont les patients qui choisissent de consulter un médecin, de suivre ses prescriptions et recommandations, ou de rechercher d'autres moyens de répondre à leurs besoins. Or, de nombreuses études ont montré qu'à l'égard de l'efficacité des traitements, les opinions des médecins diffèrent de celles des patients et de celles de leurs proches. Chacun de ces trois groupes adopte en effet des critères de jugement différents. Les médecins sont attentifs en priorité aux signes cliniques et aux symptômes ; les patients s'intéressent à ce qu'ils ressentent et à leur capacité à satisfaire leurs besoins et leurs désirs ; quant aux proches, ils accordent davantage d'importance aux comportements et attitudes face à la maladie. D'une façon générale, on peut considérer que la plupart des personnes s'orientent vers le système de santé en fonction de leur propre conception de la santé et non pas en fonction du cadre conceptuel réel de la médecine clinique.

### *Le retour à une conception plus humaniste de la relation médecin-malade*

Si la médecine moderne et son langage ont été accusés d'être impersonnels, c'est pour s'être trop éloignés des modes d'expression des patients, de leur conception de la maladie et du sens qu'elle prend pour eux. Dans la mesure où les perceptions des patients déterminent de façon cruciale l'utilisation des services et des traitements, et l'impact que ceux-ci auront sur leur état de santé, ces perceptions doivent être prises en compte directement. Les décideurs dans le domaine de la santé ont donc besoin d'indicateurs de résultats fiables et pertinents qui puissent les renseigner spécifiquement sur l'impact de leurs choix sur la qualité de vie et les perceptions des consommateurs de soins.

### *L'effet de mode*

Enfin si nous ne pouvons nier le phénomène de mode autour de ce concept, il convient de remarquer qu'il reste un concept ancien et universel, repris par de nombreux philosophes comme Socrate : *"J'attache plus de prix à la qualité de ma vie qu'au nombre de jours qu'elle a pu compter"*.

Avec les mesures de qualité de vie, la médecine a donc entrepris de changer de perspective pour adopter celle du patient ce qui l'a conduit à substituer à son objet traditionnel (la santé), la qualité de vie de ses patients. L'engouement autour de la qualité de vie, témoignant d'un véritable changement de référentiel, permet de rappeler qu'une mesure doit être dotée de propriétés métrologiques acceptées et reconnues, assurant une solution technologique alternative au jugement subjectif. Cette étape de validation doit constituer le préalable à toute appropriation par la communauté scientifique. Au-delà, l'approche développée autour du concept de qualité de vie liée à la santé a permis de rappeler que la mesure a une finalité, celle de décrire un objet caractérisé par un ou plusieurs attributs ; la mesure consistant alors en l'affectation de symboles à des objets de telle sorte que nous puissions octroyer une quantité à un attribut numérique, ou classer des objets dans une ou plusieurs catégories selon des règles de classification. La mesure consiste donc en une représentation abstraite de l'objet que nous souhaitons appréhender.

En France la prise en compte de la qualité de vie est très variable ; d'une part on a recours à une multitude d'instruments de mesure relatifs à la qualité de vie dont la validation est souvent discutable ; d'autre part il existe une réticence de la communauté médicale à revendiquer pleinement ces instruments. Parmi les causes de celle-ci, nous devons reconnaître que le monde de la santé est quelque peu désarmé : qu'est ce que la qualité de vie, à quoi cela sert-il, quels instruments sont disponibles et lesquels choisir, comment utiliser en pratique ces mesures, comment interpréter les données de qualité de vie ?

Sans pouvoir répondre complètement à l'ensemble de ces questions il est nécessaire de définir le concept de qualité de vie, d'en comprendre les utilisations et les grandes propriétés métrologiques : validité, fidélité, acceptabilité ...

## 2. Champ d'application et Utilisation

Les études dans lesquelles sont mis en oeuvre les instruments de mesure de la Qualité de Vie ont des objectifs variés. D'une façon générale, il s'agit, en quantifiant les répercussions des pathologies et des soins sur la qualité de vie des patients, de rassembler des données qui permettent de juger de la pertinence d'interventions déterminées afin d'orienter la recherche et les soins. L'hypothèse est que les mesures ainsi obtenues sont susceptibles de contribuer à expliquer la demande pour les services de santé, l'observance des traitements et la satisfaction des patients, et en définitive à contribuer à l'amélioration du bien être et de la satisfaction de la population vis-à-vis des traitements et du système de santé.

La mesure de qualité de vie peut avoir trois objectifs majeurs :

- ◆ soit la mesure a pour objectif de permettre de discriminer les personnes à un moment particulier, comme c'est le cas des tests psychologiques ou d'intelligence ; nous parlerons alors de mesure de QVLS discriminative ;

- ◆ soit la mesure a pour objectif de prédire une évolution future, et alors cette mesure se définit comme une référence ; nous parlerons alors de mesure de QVLS prédictive ;
- ◆ soit et c'est l'application la plus courante, la mesure a pour objectif de détecter les changements dans le temps, comme c'est le cas dans les études de cohorte ou les essais contrôlés ; nous parlerons alors de mesure de QVLS évaluative.

En santé publique les applications sont multiples : évaluation de la prévention, évaluation de la planification. On peut s'interroger sur les conséquences de la mise en place d'un dépistage en termes de qualité de vie ? De même l'efficacité théorique et réelle d'une action de prévention primaire est extrêmement difficile à démontrer ; la prise en compte des préférences des patients en fonction des états de santé attendus comme résultats de l'intervention permet d'éclairer le jugement des décideurs.

### 3. Validation d'une mesure de qualité de vie

La validation doit répondre à trois questions : l'instrument est-il pertinent ou valide, fiable et applicable ?

Un outil ne peut être considéré comme fonctionnel que s'il présente des qualités de mesure suffisantes.



#### La validité

La validité (ou pertinence) d'un instrument est sa capacité à bien mesurer ce qu'il est censé mesurer.

Il existe trois types de validité :

#### ◆ la validité du contenu

consiste à juger si le jeu de questions sélectionnées représente bien toutes les facettes du concept à mesurer,

#### ◆ la validité de structure interne

explore la cohérence de l'agencement interne de l'échelle, notamment par l'analyse des corrélations des réponses aux questions entre elles ou par la mise en oeuvre de techniques d'analyses de données,

#### ◆ la validité de structure externe

repose soit sur l'étude des corrélations de cette mesure avec une autre échelle de QVLS faisant référence, soit sur l'étude des liaisons avec d'autres descripteurs externes : socio-démographiques, cliniques.



#### La fidélité

Dans son sens le plus usuel, la fidélité réfère au degré d'absence d'erreur de mesure lors que nous faisons passer un test (réalisation d'une mesure) bâti sous la forme d'un score.

Dans le processus de mesure nous pouvons être confrontés à deux types d'erreurs : l'erreur systématique et l'erreur non systématique.

#### ◆ l'erreur systématique

est une erreur qui se surviendra au décours des différentes répétitions de la mesure d'un même trait exploré. Elle affecte la consistance du score, altérant le résultat toujours dans le même sens. L'erreur systématique a pour effet de donner une mesure inappropriée ; c'est donc l'utilité de la mesure qui est affectée.

#### ◆ l'erreur non systématique

est en fait l'erreur aléatoire qui varie de façon imprévisible lors des différentes mesures. Elle peut donc affecter le score tant positivement que négativement. L'erreur aléatoire affecte tant la consistance que l'utilité.

La fidélité (ou précision) de l'instrument est principalement appréciée sur deux critères :

#### ◆ la cohérence interne

appréciée par le calcul du coefficient alpha de Cronbach (il peut prendre des valeurs de 0 à 1 ; des valeurs au-delà de 0,70-0,80 sont considérées comme souhaitables),

#### ◆ la reproductibilité

qui est sa capacité à produire des scores de QVLS comparables lorsque la mesure est répétée, alors que l'état de l'individu est stable. Quand l'instrument fait appel à un évaluateur externe, l'étude de la fiabilité doit comprendre l'étude de la fidélité intercotateurs (capacité à produire des scores similaires quand à un temps donné au moins deux évaluateurs différents cotent le même sujet avec le même instrument).



#### La sensibilité au changement

La sensibilité au changement est la capacité de l'instrument à mettre en évidence une variation jugée comme pertinente par les experts.

Ce dernier critère est particulièrement requis dans un contexte d'évaluation d'interventions de santé, comme lors d'un essai thérapeutique.



#### Attention

Un instrument sera applicable si les conditions d'utilisation montrent que le coût de mise en oeuvre est modeste, que l'acceptabilité par les sujets et la communauté médicale est grande, et que le temps de passation est faible.

## Chapitre B. Théorie de la mesure

## 1. Définition

Selon une définition classique, *"mesurer consiste dans des règles d'attribution de nombres à des objets de façon à représenter des quantités d'attributs"* (5).

La réduction de l'objet à ses attributs mesurables est donc le processus fondamental de toute mesure: on ne mesure pas des objets, on ne mesure que leurs attributs.

Il ne s'agit donc pas de savoir ce qu'est la santé, la qualité de vie ou le bien-être à la manière d'Aristote ou de Platon, c'est-à-dire d'une manière spéculative. Il n'est pas question non plus de mesurer directement des entités aussi vastes que la santé, la qualité de la vie ou la satisfaction des patients. Il s'agit seulement de quantifier certains de leurs attributs caractéristiques définis dans une perspective décisionnelle : la plupart des instruments de mesure de la qualité de vie sont donc constitués de questions qui sont-elles mêmes regroupées en dimensions qui visent chacune un aspect de la qualité de vie.

## 2. Le paradoxe de la mesure de la qualité de vie

Il y a bien entendu un paradoxe à penser que l'on puisse quantifier du qualitatif et mesurer des états tels que la santé ou la qualité de vie en se basant sur des éléments d'information aussi subjectifs que les opinions des malades. Le projet d'une mesure de la qualité de vie liée à la santé est souvent critiqué au nom d'une différence irréductible qui existerait entre les "qualités" et les "quantités" (6). La plupart des controverses sur la possibilité de mesurer la qualité de vie semblent dues à une méconnaissance des limites que revêt cette entreprise. Néanmoins, la différence que l'on observe entre ce qui semble relever de l'ordre du qualitatif et ce qui semble relever de l'ordre du quantitatif, peut être considérée comme un artefact dont la cause se trouve dans notre système d'observation.

L'approche quantitative conduirait à des nombres, quand l'approche qualitative conduirait à des concepts et donc à des mots (les modalités verbales). Il s'agit en fait de deux systèmes de représentation.

Il n'est donc pas impossible de réinterpréter en termes quantitatifs des attributs initialement pensés en termes qualitatifs. L'analyse qualitative est une condition sine qua non et préalable de la mesure, puisque c'est à partir de celle-ci que les questions dont seront constitués les instruments, sont proposées. La valeur de la mesure dépend donc de celle de l'analyse qualitative.

## 3. Perspective opérationnelle



### La qualité de vie

Dans une perspective opérationnelle, la qualité de vie doit donc être définie comme un ensemble d'attributs mesurables pourvu d'au moins un élément.



### Démarche

La plupart de ses instruments de mesure sont des ensembles de questions standardisées qui sont regroupées en dimensions ou "concepts" mesurés. On interroge successivement, par exemple l'activité physique, l'état psychologique, l'activité sociale et le sentiment de bien-être. Des scores sont attribués aux différentes réponses possibles. Un algorithme permet de calculer le score qui correspond à chaque dimension (et éventuellement le score global).

En ce qui concerne les mesures de la qualité de vie, cette représentation quantitative est tributaire de la représentation qualitative préalable d'une qualité déterminée (telle qu'une difficulté à se mouvoir) à laquelle sera ensuite attribuée une grandeur qui prendra la forme d'un score. Il est clair que le projet de mesurer la qualité de vie n'exclut nullement l'approche en termes qualitatifs du réel, tout au contraire.

La quantification de la qualité de vie a seulement pour objectif de développer une connaissance moins ambitieuse, mais plus efficace peut-être que certaines analyses qualitatives. Sa seule légitimité ne peut donc se trouver que dans la considération de ses effets : il faut et il suffit qu'elle autorise des interventions plus efficaces, qui permettent d'améliorer la santé et le bien-être des individus et des populations

Ainsi, les mesures de qualité de vie reposent, au moins pour une part, sur les jugements que portent les sujets eux-mêmes sur leur état de santé. A ces sujets, on demande de dire dans quelle mesure ils sont satisfaits de leur santé ou de la vie qu'ils mènent, et dans quelle mesure leur existence a été modifiée par leur maladie et par les interventions médicales qu'ils ont subies.

Les mesures de Qualité de Vie liée à la santé sont obtenues à partir de l'analyse des réponses des sujets à un questionnaire standardisé, mais le questionnaire n'est pas l'instrument (7).

Les trois composantes constitutives de tout instrument de mesure de la qualité de vie sont :

#### **1. un ensemble de "descripteurs"**

figurant dans le questionnaire qui vont permettre de définir différents états de santé auxquels chaque sujet est identifié. Les questions peuvent être regroupées en dimensions (ou concepts mesurés) dont on pense qu'ils sont liés à la qualité de vie : par exemple, la mobilité, la vie et les relations avec les autres.

#### **2. un ensemble de pondérations ou valeurs subjectives**

qui sont associées à chacun des états possibles et les situent quantitativement

les uns par rapport aux autres. Ces valeurs subjectives sont censées refléter la hiérarchie implicite par rapport aux états de santé du groupe de population concerné. Dans la perspective de l'utilisation de ces mesures pour informer les décisions de santé publique, l'ensemble de ces valeurs peut-être assimilé à ce que les économistes appellent une fonction de choix social.

### 3. la fonction d'attribution des valeurs à chacun des états possibles,

c'est-à-dire la formule, ou " algorithme de calcul des scores ", qui permet de calculer chaque score et qui fixe les relations entre les réponses des sujets aux questions et la mesure des différents concepts interrogés.

En conclusion, les instruments sont constitués de questions (ou items), qui peuvent être regroupés en dimensions ou concepts, auxquels est associé un système de recueil des réponses des sujets, et d'un algorithme qui permet d'attribuer une pondération aux différentes réponses et de calculer les scores. Ces mesures qui associent donc des valorisations subjectives à des éléments descriptifs quasi objectifs sont appelées, pour cette raison, des mesures composites.



### Exemple

Question 9a, 9b, 9d : Les questions qui suivent portent sur comment vous vous êtes senti(e), AU COURS DE CES 4 DERNIERES SEMAINES. Pour chaque question, merci d'indiquer la réponse qui vous semble la plus appropriée. Au cours de ces 4 DERNIERES SEMAINES, y a-t-il eu des moments où : (entourez la réponse de votre choix, une par ligne).

	En permanence	Très souvent	Souvent	Quelquefois	Rarement	Jamais
a. Vous vous êtes senti(e) dynamique ?	1	2	3	4	5	6
b. Vous vous êtes senti(e) très nerveux (se) ?	1	2	3	4	5	6
c. Vous vous êtes calme et détendu(e) ?	1	2	3	4	5	6

↑ TAB. 1 : QUESTIONNAIRE SF-36

## Chapitre C. Les concepts mesurés

Pour répondre à ce besoin, psychologues, économistes et chercheurs en santé publique ont mis au point, dès les années soixante-dix, des instruments et des méthodes qui permettent de mesurer les effets des interventions de santé en termes de statut de santé, de qualité de vie ou de préférences liées à la santé.

Les instruments n'ont pas pour ambition de mesurer directement des entités comme la santé, la qualité de vie ou la satisfaction des patients, mais simplement de quantifier certains de leurs attributs caractéristiques identifiés grâce aux définitions opérationnelles.

Chaque définition opérationnelle est construite en fonction des objectifs que l'on veut atteindre. Ces objectifs sont, par-delà la simple quantification de la répercussion de la pathologie, de produire des informations contrôlées, reproductibles et fiables, destinées à orienter la recherche, les soins, à informer les décideurs et à améliorer le bien-être et la satisfaction de la population vis-à-vis des traitements ainsi que vis-à-vis du système de santé.

Les principaux concepts visés par ces définitions sont la santé, la qualité de vie liée à la santé et la qualité de vie.

## 1. La santé

Les premières mesures utilisées dans le domaine de la santé ont été élaborées pour quantifier la santé des sujets ; ainsi, un certain nombre d'instruments valides et fiables ont été développés aux États-Unis ou au Royaume-Uni pour mesurer, par exemple, la souffrance perçue, l'impact de la maladie (8), le fonctionnement physique et la satisfaction avec la vie (9). Il y avait aussi un certain nombre d'instruments destinés à mesurer l'état de santé dont les questions portaient sur des symptômes physiques et psychologiques (10).

## 2. La qualité de vie liée à la santé

Quand la demande pour des mesures de "qualité de vie" est apparue, de telles mesures n'étaient pas disponibles car jusqu'alors les recherches avaient porté sur la mesure de l'état de santé. Le terme de "qualité de vie liée à la santé" fut alors inventé pour justifier l'utilisation des mesures disponibles en changeant leur nom. Le raisonnement tenu était le suivant : puisque ces questionnaires sont centrés sur les aspects de l'existence qui sont affectés par une mauvaise santé, ils doivent aussi nous donner des indications sur l'impact de la maladie sur la qualité de la vie. Le concept de "qualité de vie liée à la santé" implique que l'on puisse analyser la qualité de la vie dans ses composantes "liées à la santé" et ses composantes "non liées à la santé". Cette perspective ne tient pas compte des liens du statut de l'état de santé avec les autres aspects de l'existence, comme : les habitudes de vie, le contexte culturel, les croyances religieuses ou non, le statut professionnel, les revenus...

Pour Patrick et Erickson (11) : *"la qualité de vie liée à la santé est la valeur qui est attribuée à la durée de la vie en fonction des handicaps, du niveau fonctionnel, des perceptions et des opportunités sociales modifiées par la maladie, les blessures, les traitements ou les politiques de santé"*.

Pour M. Bergner (12), les principaux domaines de la Qualité de Vie sont :

- ◆ les symptômes,
- ◆ le statut fonctionnel,
- ◆ les activités liées au rôle social,
- ◆ le fonctionnement social,
- ◆ la cognition,
- ◆ le sommeil et le repos,
- ◆ l'énergie et la vitalité,
- ◆ l'état ou le statut émotionnel,
- ◆ la perception de la santé et enfin
- ◆ la satisfaction générale à l'égard de la vie.

Pour une majorité d'auteurs, la Qualité de vie liée à la santé (QVLS) présente un caractère multidimensionnel. L'identification de ces dimensions et l'importance respective (la pondération) qui leur est attribuée lors de la mesure sont des enjeux importants. Cependant il existe aujourd'hui un consensus minimal pour dire qu'une mesure de qualité de vie doit conceptuellement prendre en compte trois dimensions : physique, psychologique, social et considérer éventuellement des dimensions supplémentaires plus ou moins spécifiques d'un groupe de patients ou d'une pathologie. Ainsi à ces trois dimensions certains auteurs ajoutent la notion de sensations somatiques (symptômes, conséquences des traumatismes ou des procédures thérapeutiques, douleurs), ce qui est discutable en tant qu'individualisation d'une dimension. Concernant la dimension sociale, on peut vouloir représenter des notions comme la participation aux événements sociaux, la qualité des relations amicales, familiales et sociales, la satisfaction de la vie maritale, la satisfaction professionnelle ... Concernant la dimension physique, les aspects de la vie quotidienne doivent être pris en compte, tout comme la symptomatologie clinique (douleur, fatigue), les loisirs, le travail. Concernant la dimension psychologique seront particulièrement relevés, l'humeur (dépression, anxiété), les performances des fonctions cognitives (mémoire, concentration), le bien-être. Le contenu de ces dimensions est fonction de la conceptualisation de la qualité, faite autour de la santé. Les choix doivent toujours être justifiés et décrits afin d'aider le futur utilisateur des mesures ainsi constituées.

### 3. La qualité de vie

Les instruments actuels de QVLS peuvent apparaître comme ne prenant pas suffisamment en compte la complexité de l'expérience humaine, qui reste pourtant le centre d'intérêt principal des médecins et des patients (13). A ce type de démarche doit être opposée une autre approche, plus soucieuse de l'expérience des sujets, qui nous aide à comprendre comment une réalité apparemment insupportable peut parfois être une source positive de bien-être et de qualité de vie. La qualité de vie est un concept équivoque comme peut l'être celui de l'intelligence. Dans ces deux cas ces concepts sont dotés d'un sens commun mais variable d'un individu à l'autre et difficilement communicable. La qualité de vie peut être affectée de significations différentes reflétant l'expérience, la connaissance et les valeurs individuelles. Elle peut donc prendre en compte des significations aussi large que la culture, la philosophie, la vie relationnelle, la vie spirituelle, les finances, la politique, la psychologie, la santé. Le concept de QV accepte plusieurs définitions qui sont standardisées et explicites, mais qui sont multiples : les définitions varient selon les écoles et selon les objectifs de la mesure.

Selon Nordenfelt (14), il s'agit, pour évaluer la qualité de vie, de considérer :

- ◆ l'ensemble des perceptions, des sensations, des émotions et des idées d'une personne,
- ◆ l'ensemble de ses actions et de ses activités,
- ◆ l'ensemble de ses réussites,
- ◆ l'ensemble des événements et des actions qui l'affectent. L'OMS, dans une perspective de santé publique, définit la qualité de vie comme *"la perception qu'a un individu de sa place dans l'existence, dans le contexte de la culture et du système de valeurs dans lesquels il vit, en relation avec ses objectifs, ses attentes, ses normes et ses inquiétudes (15)." Le fait de mettre l'accent sur la culture est caractéristique du projet OMS, qui vise à développer une mesure valable pour l'ensemble du monde.*

La qualité de vie peut également être définie par rapport aux besoins perçus par les patients. Quels sont ces besoins et ces désirs qu'il conviendrait d'interroger pour évaluer la qualité de vie? Selon la conception, soutenue par Hunt et Mc Kenna (16), chaque situation pathologique interfère de façon particulière avec la capacité des individus à satisfaire leurs besoins. Si les besoins perçus sont satisfaits, la qualité de vie est augmentée ; s'ils sont peu satisfaits, la qualité de vie est diminuée. Le point clé est alors de comprendre l'importance des besoins perçus par les patients, et non pas seulement les besoins identifiés par les médecins. Ainsi les besoins suivants sont des besoins fréquemment rapportés comme devant être satisfaits afin de présenter des niveaux de qualité de vie élevée :

- ◆ alimentation, sommeil, activité, vie sexuelle, absence de douleur
- ◆ logement, sécurité, stabilité
- ◆ affection, amour, communication, sentiment de communauté, de filiation

- ◆ curiosité, jeux, créativité, sentiment que sa vie a un sens
- ◆ identité, reconnaissance sociale, respect, estime de soi, capacité de se réaliser.

## Chapitre D. La question du point de vue

### 1. Définition

Selon le Littré, le point de vue c'est, au sens propre, "l'endroit où est placé celui qui regarde un objet". En d'autres termes, la réduction de l'objet à ses dimensions "observables" mesurables suppose l'existence un point de vue à partir duquel elle est opérée. Choisir un point de vue est une nécessité scientifique, c'est la condition de possibilité de toute représentation et de toute mesure. Toute mesure de la santé et de la qualité de vie suppose donc l'existence d'un point de vue à partir duquel elle s'opère.

Ensuite, toujours selon le Littré, dans un sens plus spécialisé, le point de vue, c'est "l'endroit où l'on doit se placer pour voir un objet, une scène, le mieux possible". D'une part, cette seconde définition confirme, l'idée d'une multiplicité de points de vue sous lesquels peut être observé un même objet : selon le point de vue adopté par l'observateur l'apparence de l'objet sera différente. D'autre part, elle introduit l'idée que cette multiplicité peut être hiérarchisée -tous les points de vue ne sont pas équivalents ce qui implique que la valeur de l'observation dépend de la façon dont l'observateur se place. Bien entendu, on notera que différents points de vue peuvent se compléter et permettent ensemble de considérer, par exemple une situation médicale et ses enjeux dans leur globalité. Mais, il reste que tout observateur a pour devoir de se placer au point de vue qui permet de se représenter l'objet "le mieux possible".

### 2. Difficultés conceptuelles

L'analyse du point de vue permet donc de préciser "qui" considère l'objet à étudier et "comment" il se met en rapport avec cet objet, à quelle distance, sous quel angle, selon quel éclairage. Il s'agit donc d'une notion essentielle pour comprendre le projet de mesure de la santé et de la qualité de vie liée à la santé. La complexité conceptuelle du projet d'une mesure de la qualité de vie vient de ce que le sens figuré de "point de vue" -celui d'une opinion parmi d'autres possibles-, interfère avec celui de ses sens propres qui admet qu'il existe un point de vue où l'observateur doit se placer pour représenter le réel le plus objectivement possible. Ensuite, dans le cas de la mesure de la qualité de vie, nous sommes toujours dans une situation de double point de vue : le point de vue de celui qui considère (par exemple, le médecin) et le point de vue de celui qui est considéré (le patient). Le médecin examine le malade, le malade est vu sous le point de vue du médecin, mais cet objet de la pensée du médecin est aussi, de son côté, un sujet qui a donc aussi son propre point de vue. Le premier problème qui se pose est donc de savoir si, pour mieux soigner le malade, le médecin ne devrait pas essayer (afin de compléter son point de vue médical) de le voir selon le point de vue que le malade a sur lui-même ? Dès lors se pose le problème de savoir comment construire des indicateurs de qualité de vie liée à la santé qui reflètent réellement le point de vue du patient ?

### 3. Les différents points de vue possibles

La définition de la qualité de vie (et des besoins) des patients par les professionnels de la santé semble contestable car ces derniers ont tendance à les rapporter aux techniques particulières dont ils savent user avec compétence. Un besoin de mobilité peut, par exemple, être identifié, par les professionnels, à un besoin de prothèse totale de hanche, ou de prescription d'anti-inflammatoire, etc. Cette tendance gêne la mise en relation adéquate de besoins des sujets et de leurs désirs variés avec l'ensemble des moyens qui permettraient de les satisfaire. Afin de contourner cette difficulté, on a entrepris de développer des questionnaires destinés à estimer, le plus directement possible, la qualité de vie des sujets telle qu'ils la ressentent eux-mêmes. Dans cette perspective, l'estimation des besoins ressentis consiste à comparer la situation présente des sujets à une liste de besoins minima, aux besoins moyens de la population à laquelle appartiennent les sujets, voire à une référence idéale (par exemple un état de santé parfait). La définition de la santé par l'OMS est un exemple de ce type de référence. Mais est-il vraiment possible de définir un idéal qui soit universel ?

Afin de mettre au point des indicateurs qui contribuent utilement à l'élaboration des politiques publiques, des listes de besoins minimaux dont on considère qu'ils doivent être impérativement satisfaits ont été proposés. Dans le secteur de la santé, les instruments construits selon ce cadre conceptuel visent à mesurer le degré selon lequel les sujets sont capables du fonctionnement minimal de tout adulte en bonne santé (marcher, s'habiller, travailler, etc.). Mais cette approche tend à mettre de côté de nombreux besoins qui peuvent être également légitimes (le besoin de reconnaissance sociale par exemple).

Les besoins peuvent également être estimés par rapport à une population de référence. Cette population peut être comparable à la population à laquelle appartiennent les sujets dont on évalue la qualité de vie, mais il est aussi possible de prendre comme référence le groupe des membres les plus privilégiés de la société.

Enfin, certains instruments visent à refléter la façon dont les individus concernés apprécient eux-mêmes l'écart qui existe entre leur situation présente et une situation hypothétique où leurs besoins seraient satisfaits. Ces instruments sont adaptés à l'évaluation des soins de santé primaires dans la mesure où c'est toujours la perception d'un écart qui conduit les sujets à consulter un médecin. Cette approche a l'inconvénient de dépendre des espérances particulières des sujets. On a donc proposé de déterminer les besoins de référence en interrogeant la population dans son ensemble. Les besoins d'un individu sont alors estimés par rapport à des standards jugés acceptables par les autres. Mais la moyenne des avis d'un groupe de sujets peut n'avoir que peu de rapports avec l'expérience de chaque individu. Cette dernière approche, pour être préférable, n'est donc pas pour autant parfaite.

Par conséquent, lorsque l'on choisit un instrument, on doit toujours se demander comment ont été définis les besoins dont la satisfaction est évaluée et qui a procédé à cette définition. En conclusion, pour identifier le point de vue caractéristique d'un instrument donné, il faut examiner chacune des trois composantes structurelles de toute mesure. C'est-à-dire qu'il faut examiner les modalités concrètes du développement, de la validation, de l'adaptation et de l'utilisation des instruments afin de pouvoir répondre aux questions suivantes :

- ◆ les questions de qui ? (quel est l'origine et le mode d'identification des questions),
- ◆ les valeurs de qui ? (quel est l'origine et le mode de calcul des valeurs),
- ◆ les théories de qui ? (analyse critique des théories).

Ainsi, l'analyse du concept de point de vue contribue à la description de tout instrument de mesure. A ce titre, elle s'insère dans la panoplie des critères qui permettent de sélectionner un instrument pour un usage donné. Par exemple, la caractérisation du point de vue d'une mesure peut, associée à d'autres considérations plus méthodologiques, contribuer à évaluer la possibilité de généraliser des résultats obtenus dans une population donnée.

## Chapitre E. Typologies des instruments de mesure de qualité de vie

Les instruments peuvent être classés en fonction de chacun de ces éléments constitutifs.

### 1. En fonction du champ d'application

On oppose classiquement les instruments génériques aux instruments spécifiques (exemples 1 et 2).

**◆ Les instruments génériques**

sont élaborés à partir de questionnaires mis au point en population générale. Ils ne peuvent jamais rendre compte avec détail de la spécificité de chaque pathologie abordée. De ce fait les instruments génériques sont en général peu sensibles aux variations d'état de santé. Cependant, ils permettent de comparer les résultats à ceux d'autres études menées sur des groupes de patients différents, sur des pathologies différentes, ou dans des pays différents.

**◆ Les instruments spécifiques**

sont eux adaptés à un problème particulier. L'instrument peut être spécifique d'une pathologie donnée (cancérologie, asthme ...) ou d'une population donnée (enfants, population âgée ...). On pourra choisir, seul ou en association, l'un ou l'autre de ces types d'instruments selon l'objectif poursuivi.

**Exemple : SF-36 : Short-Form 36**

Le SF-36 est un instrument nord-américain générique, s'adressant à la population générale de plus de 14 ans, de qualité de vie. Il permet d'évaluer l'état de santé dans la pratique et la recherche clinique. Il est aussi utilisé pour l'évaluation des politiques de santé. Il est rempli par le sujet lui-même, comporte 36 questions, explorant 8 domaines principaux au cours des quatre dernières semaines : Activité physique, limitations dues à l'état physique, douleur, vie et relations avec les autres, santé psychique, limitations dues à l'état psychique, vitalité, santé perçue. Un score de qualité de vie peut être calculé simplement pour chacun de ces huit domaines. Cet instrument est largement diffusé dans plusieurs pays et plusieurs langues comme l'anglais, l'allemand, le français, l'espagnol, le finnois, le danois, le portugais, le polonais, le russe ... permettant des comparaisons internationales. Cependant d'autres instruments génériques sont disponibles comme le Duke Health Profile, l'Euroqol, le Sickness Impact Profile ...

**Exemple : Le Childhood Asthma Questionnaire : un instrument spécifique de l'enfant asthmatique**

L'asthme de l'enfant est une pathologie posant un grave problème de Santé Publique. En effet sa prévalence est élevée, sa mortalité est inquiétante, sa morbidité associée est préoccupante. Les instruments de qualité de vie liés à la santé développés dans l'asthme, l'ont été principalement chez l'adulte. Si pendant longtemps la qualité de vie chez l'enfant asthmatique, n'a été appréhendée que par l'intermédiaire d'une tierce personne, médecin, parent ou enseignant, nous disposons maintenant en France d'un autoquestionnaire de qualité de vie lié à la santé développé et validé en Grande-Bretagne et en France s'adressant directement aux enfants asthmatiques de 4 à 16 ans : le Childhood Asthma Questionnaire. Pour les enfants de 7 à 11 ans, 23 questions sont remplies sans l'aide de parents. Cet outil permet d'une part d'apprécier les répercussions sur l'activité quotidienne commune à tous les enfants (sports, jeux, lecture, dessin, télévision...), définissant deux dimensions la Qualité de Vie Active et la Qualité de Vie Passive, d'autre part d'explorer plus spécifiquement l'asthme (fréquence des crises, gravité, attitude par rapport au traitement...) au travers de deux axes la détresse et la sévérité. Ce questionnaire permet d'objectiver des bénéfices personnels, non repérés par les jugements para cliniques classiquement utilisés comme le volume expiratoire, le débit expiratoire de pointe.

## 2. En fonction de l'origine des questions

Les questions et les systèmes de recueil des réponses peuvent avoir plusieurs sources :

- ◆ les experts (par exemple les psychométriciens qui ont développé l'instrument),
- ◆ les groupes de patients,
- ◆ les médecins,
- ◆ d'autres travailleurs sociaux ou médicaux,
- ◆ la famille,
- ◆ un tiers.

L'appréciation par l'un quelconque de ces partenaires devrait être conditionnelle au point de vue, que conceptuellement l'échelle souhaite explorer : est-ce la qualité ressentie par le patient, est-ce la qualité objectivée par le médecin ... ? L'appréciation par le médecin présente l'avantage de pouvoir apprécier la qualité de vie de pratiquement tous les patients, dont notamment ceux qui n'auraient pas pu remplir un auto-questionnaire (sujets handicapés..). Cette méthode très consommatrice de "temps médical", reste rapide. Elle présente le gros avantage de diminuer considérablement les réponses manquantes (questions difficiles à comprendre pour le patient...), mais se pose le problème de la cohérence entre différents "investigateurs cotuteurs".



### Conseil

Les auto-questionnaires sont en général plus pratiques, moins coûteux, plus simples (permettant de les administrer plusieurs fois pendant l'essai). Ils permettent toujours en cas d'échec de passer à l'interrogatoire par le médecin sous réserve de respecter certaines conditions. La critique essentielle tient dans le nombre plus important de données manquantes pouvant être pour la plupart évitées par des explications claires préalables à la première administration. De notre point de vue dès lors que nous parlons de santé perçue par le patient il convient le plus souvent possible de recourir à l'auto-questionnaire.

### 3. En fonction du type de mesure obtenue

Suivant les scores que l'on peut obtenir, on distingue deux types d'instruments (17) :

- ◆ les index qui permettent de calculer un score global et,
- ◆ les profils qui fournissent un score pour chacune des différentes dimensions explorées de la qualité de vie, sans les combiner en un score unique.

Au final plusieurs méthodes permettent d'obtenir les pondérations. Les scores peuvent être calculés simplement en additionnant la valeur des réponses des sujets à chacune des questions.



#### Exemple

Par exemple cette valeur peut être 1 si la réponse est oui et 0 si la réponse est non. Dans un tel cas, on obtient une pondération de l'importance des attributs par le nombre de questions qui les concernent, chacune d'entre elles se trouvant dotée d'une pondération égale.

On peut également se trouver confronté à une somme pondérée par les utilités. Les pondérations sont alors issues soit de l'étude des préférences des sujets interrogés, soit de l'étude des préférences d'un groupe de juges représentant les patients ou de tout autre groupe de la société. Ces pondérations, qui ont la faveur des économistes, sont relativement complexes à obtenir (18). Enfin, les pondérations sont parfois obtenues par technique statistique. Dans tous les cas, la méthode utilisée et la composition de la population de référence doivent être précisées car elles peuvent influencer les pondérations résultantes.

\* \*  
\*

Malgré les efforts de classification tant théorique que conceptuelle, la définition de la qualité de vie liée à la santé reste une approche complexe où chacun emploie sa propre terminologie, exprimant par la même son propre point de vue. Il n'existe donc pas aujourd'hui de définition claire, univoque, simple et encore moins universelle de la QVLS. Il ne serait d'ailleurs pas licite de vouloir contraindre les concepteurs de mesure de qualité de vie, dans le cadre d'une définition spécifique au risque de stériliser l'évolution de notre pensée. Cependant l'ensemble des approches conceptuelles tend à montrer que dans le domaine de la QVLS, il existe une volonté partagée de quantifier en termes scientifiquement analysables les conséquences d'une maladie et de son traitement sur la perception individuelle du patient de sa capacité à vivre une vie utile et entière. De cette définition nous retiendrons que la QVLS est donc un concept mesurable, évaluable dans le temps, basé sur une approche subjective dont les applications dans le champ de la santé publique sont multiples.

## Bibliographie

1. Leplège A. La mesure de la qualité de vie, Coll Sue sais-je ? PUF, 1999, n° 3506.
2. Hunt S.M. Cross-cultural comparability of measures and other issues related to multicountry studies. *Brit J Med Econ* 6C 27-34, 1993.
3. Jecker N.S. The Medical Futility Debate in the United States, Premier Colloque Mondial Philosophie et Médecine, Paris, 1994.
4. WHO Chron. 1947.
5. Nunnally J. Psychometric theory (2nd ed). New York McGraw-Hill ; 1978.
6. Furtos J. Qualité de vie règne de la quantité. *L'information Psychiatrique*. 5, XVII-XX, 1992.
7. Barofski I., Sugarbaker P, Cancer, in Spilker B. (Ed) Quality of life assessments in clinical trials, Raven Press, New York ; 1989.
8. Bergner M., Bobitt R., Pollard W., et al. The Sickness Impact Profile : development and final revision of a health status measure. *Medical Care*. 1976.
9. Stewart A , Ware JE Jr (Ed) Measuring Functioning and Well-Being - The Medical Outcome Study Approach, Duke University Press, Durham NC, 1992.
10. Elinson, J, Siegmann AE Socio-medical Health Indicators. Baywood Publishing Co. Inc. Farmingdale, New York, 1976.
11. Patrick DL, Erickson P. Health status and Health Policy, Quality of life in health care evaluation and resource allocation. Oxford University, New York, 1993.
12. Bergner M. Quality of life, health status and clinical research. *Medical Care*. 1989, 27 :S148-156.
13. Hodge J. The quality of life : a contrast between utilitarian and existential approaches. In Baldwin S et al (eds) Quality of life ; Perspectives and Policies 1990.

14. Nordenfelt, in Nordenfelt ed. *Concept and Measurement of Quality of Life in Health Care*, 1994.
15. World Health Organisation Working Group, 1994.
16. Hunt SM, McKenna SP. The QLDS. A scale for the measurement of quality of life in depression. *Health Policy* 1992 ; 22 :307-319.
17. Bullinger M. Indices versus profiles - advantages and disadvantages. In Walker SR et Rosser RM (eds.) *Quality of life assessment, Key Issues in the 1990's*, Kluwer Academic Publishers, Dordrecht, 1993.
18. Kind P. The development of health indices, in Teeling Smith G. (Ed) *Measuring Health : A Pratical Approach*. John Wiley and Sons, 1988.

# Développement d'un instrument de mesure de qualité de vie

## Préambule

Nous pouvons au travers de la littérature définir une approche systématisée de construction de cette mesure. Nous allons ici exposer le contenu des premières étapes intrinsèques à la production initiale d'un test de mesure. Cependant nous insisterons avant tout sur l'absolue nécessité de préciser et d'objectiver les moyens nécessaires à la construction d'une mesure. Ce point précis ne participe pas directement de la construction d'une mesure mais constitue le préalable à toute mise en place de procédure de construction.



### Démarche

Crocker et Algina (1) identifient classiquement 10 étapes :

1. Identifier le propos premier pour lequel le test est développé,
2. Identifier les comportements qui représentent les domaines dans lesquels nous souhaitons travailler,
3. Préparer la spécification du test, en approximant le nombre de questions qui doivent être appréhendées dans chacune des dimensions isolées,
4. Construire le jeu de questions à tester,

5. Réviser et discuter les questions construites,
6. Sélectionner les questions après un tri en les testant sur un petit échantillon de personnes,
7. Faire passer les questions retenues sur un large échantillon de personnes,
8. Déterminer les propriétés statistiques du score, et éliminer les questions ne répondant pas aux objectifs poursuivis,
9. Mener les études de validité et de fidélité sur le jeu de données,
10. Développer des recommandations pour les futurs utilisateurs spécifiant : les modalités d'administration, le type de scorage, les propriétés principales de l'échelle ....

Les six premières étapes sont préliminaires au premier passage sur l'échantillon de sujets explorés, les trois dernières sont secondaires aux enseignements fournis par l'expérimentation du test sur une large population.



### Complément

Cette liste est une liste minimale qu'il convient de suivre lorsque nous souhaitons développer des tests qui soient des mesures utiles et pertinentes. Les concepteurs d'échelle de mesure peuvent ajouter un certain nombre d'étapes, ou répéter de façon itérative un certain nombre de ces étapes.

Juniper et coll. (2) proposent de respecter le plan suivant relatif au développement d'un instrument de mesure : Spécification des objectifs, génération des items, réduction des items, mise en forme du questionnaire, prétest, étude de la fidélité, étude de la sensibilité au changement, étude de la validité, interprétabilité.

Il n'y a pas en l'occurrence de règle absolue, si ce n'est de respecter le plan initialement rapporté.



### Conseil

Nous remarquons que ces règles qui valent pour la construction cohérente d'une mesure, sont aussi sensiblement celles auxquelles nous pouvons nous référer lors de la sélection d'un instrument de mesure. En effet nous savons que les contraintes liées à la conception et à la validation d'un instrument de mesure sont lourdes, et que le temps nécessaire à la mise au point, à la validation et à la publication de celui-ci peut prendre de 5 à 10 ans parfois. Aussi l'utilisateur lambda aura tout intérêt à recourir à un outil déjà existant dès lors qu'il souhaite explorer la qualité de vie liée à la santé dans le contexte qui le préoccupe.

Il semble alors licite de proposer quelques grandes règles à respecter dans la sélection d'un instrument. Quelques auteurs se sont essayés à l'élaboration de recommandations dans ces circonstances (2, 3).

Nous pouvons résumer leurs recommandations à trois étapes essentielles :

- ◆ Préciser les objectifs de la mesure à opérer,
- ◆ Reconnaître les qualités métrologiques de l'instrument,
- ◆ Définir la réalisation pratique.

Il est absolument nécessaire de préciser avant la construction mais aussi à la sélection de l'outil les moyens dont on dispose. Doivent être particulièrement appréciés les moyens humains, financiers et le délai avant lequel on souhaite avoir une réponse. La sélection de l'instrument, la réalisation des pré-études et des études, l'analyse sont extrêmement consommatrices de temps. La simple élaboration du questionnaire peut nécessiter 10 à 20 versions avant d'être prétestée (3). Il est indispensable d'établir un planning des différentes opérations afin d'évaluer le temps nécessaire ainsi que les moyens matériels et humains dont il faudra disposer. La connaissance de ces contraintes fixe clairement les limites de la mise en oeuvre de ce type d'étude.

## Chapitre A. L'objectif du développement de la mesure

### 1. Généralités

Il convient de repérer clairement le propos pour lequel on souhaite développer la mesure. En effet dans le cadre de la production d'une décision médicale, il faut alors :

- ◆ identifier les différentes alternatives possibles dans le problème qui nous préoccupe,
- ◆ identifier la (les) population (s) intéressée (s),
- ◆ préciser les objectifs de la mesure.

La définition des objectifs est donc capitale. Elle permet d'une part au concepteur de développer une démarche cohérente, d'autre part, à l'utilisateur de reconnaître si l'instrument est adapté à son problème et aux populations auxquelles il s'adresse.

### 2. Définitions

#### *Mesurer pourquoi ?*

Nous pouvons identifier trois objectifs majeurs (4)

- ◆ Soit la mesure a pour objectif de permettre de discriminer les personnes à un moment particulier, comme c'est le cas des tests psychologiques ou d'intelligence ; nous parlerons alors de mesure de QVLS discriminative,

- ◆ Soit la mesure a pour objectif de prédire une évolution future, et alors cette mesure se définit comme une référence ; nous parlerons alors de mesure de QVLS prédictive,
- ◆ Soit et c'est l'application la plus courante, la mesure a pour objectif de détecter les changements dans le temps, comme c'est le cas dans les études de cohorte ou les essais contrôlés ; nous parlerons alors de mesure de QVLS évaluative.

### *Qui mesure-t-on ?*

Il convient alors de préciser si nous souhaitons obtenir une mesure de l'état de santé de la population en général ou une mesure des états de santé particuliers. Dans le premier cas il conviendra de recourir à des indicateurs génériques alors que dans le second se seront des indicateurs spécifiques.

### *Que mesure-t-on ?*

Il faudra préciser si l'objet de la mesure est de quantifier les états de santé afin de les ordonner, ou si l'objectif à moyen terme sera de définir des préférences des patients quant à ces états de santé.

### *Comment mesure-t-on ?*

Le choix d'administrer un instrument standardisé ou non est souvent débattu. Pour notre part, dans la mesure où par définition un instrument non standardisé est un instrument nouveau développé dans une nouvelle application, la question du choix ne se pose pas puisque seul l'instrument standardisé sera validé. Par contre, en termes de développement il est alors évident que le recours à des instruments non standardisés puisse apporter des réponses pertinentes. Cette question n'a en fait de sens profond que pour la construction d'index ou de profil. L'utilisation de batterie permet d'ailleurs à ce titre de répondre à une réponse spécifique en sélectionnant les instruments les plus pertinents, aux limites près de ces propriétés.

## **3. Discussion**

Selon la définition de ces objectifs, la sélection du type d'instrument peut varier.

Par exemple, une échelle de qualité de vie générique liée à la santé discriminera facilement des populations en bonne santé de celles malades à l'aide d'un nombre réduit de questions appréhendant des "traits" relativement fixes (immuable).

Inversement ce même questionnaire serait inefficace pour évaluer un changement dans le temps de cette même population malade, où il conviendrait d'explorer les comportements les plus susceptibles de varier (4).

En pratique nombre d'outils actuellement disponibles ont été développés dans un souci de discrimination de population et non pour explorer les variations dans le temps. Cette propriété de sensibilité au changement, est d'ailleurs dans ces cas très rarement documentée.

Bien que certains instruments prétendent répondre à ces trois objectifs, il reste relativement difficile pour eux d'atteindre un niveau d'efficacité maximal dans ces trois domaines (2).

## Chapitre B. Le contenu

Classiquement, le groupe chargé du développement du test va conceptualiser le trait qu'il souhaite mesurer. Il tentera à partir de la formulation de questions qu'il pense être en rapport avec ces dimensions conceptualisées de se rapprocher le plus possible des comportements qu'il explore. Cette approche n'est pas sans inconvénients : représentations de domaines sans signification pour les sujets, omissions de certaines aires par le concepteur. Il n'y a pas de définition objective dans la façon d'aborder le construit. Cependant Crocker et Algina (1) proposent certains supports susceptibles de renforcer la validité de la conceptualisation. Ainsi cette conceptualisation peut être faite au regard de différents points de vue qui ne sont pas exclusives et dont l'emploi est assujéti aux moyens financiers dont dispose le concepteur.

Il est recommandé de documenter le construit au moyen d'une des démarches suivantes.

### 1. Analyse de contenu

L'analyse de contenu consiste en la mise en place d'un entretien avec les sujets, objet de la mesure, autour du thème à appréhender. L'analyse du contenu des réponses permet d'isoler quels peuvent être les principaux axes structurant le trait à apprécier. A partir de ces catégories nous construirons des questions en rapport avec celles-ci. L'observation directe des sujets peut dans certains domaines permettre d'identifier les différentes dimensions structurant le trait à mesurer.

### 2. Analyse de la littérature

La recherche dans la littérature peut permettre d'isoler les principaux comportements rapportés par les scientifiques, relatifs au trait exploré. Cette méthode présente les avantages et les inconvénients classiquement décrits dans la revue de la littérature selon la façon dont elle est conduite et synthétisée notamment. Cependant au décours de celle-ci on peut vouloir se rattacher à une école de pensée donnée qui décrit le phénomène suivant différents axes.

### 3. Jugement d'experts

Les experts peuvent être interrogés directement sur la façon dont ils conceptualisent eux-mêmes le problème. Cette méthode est l'une des plus fréquemment utilisées pour le développement d'indicateurs dans le domaine de la santé. Les problèmes tiennent dans la représentativité de ces experts. D'autre part, dans le cadre de construction d'indicateurs de santé perceptuelle, nous pouvons nous interroger sur la légitimité de ceux-ci à définir le cadre dans lequel le patient doit exprimer son point de vue.

## Chapitre C. Construction du questionnaire

### 1. Construction des questions



#### Démarche

Le choix du type de construction de questions est extrêmement important dans la mesure où il répond à la question comment mesurer ? Pour développer les questions nous devons suivre un certain nombre de règles de construction simple:

- ◆ Sélectionner le format approprié des énoncés,
- ◆ Vérifier que le format choisi est adapté aux sujets qui seront évalués par l'instrument de mesure,
- ◆ Sélectionner les rédacteurs des questions,
- ◆ Écrire les questions,
- ◆ Explorer la qualité des questions.

#### Sélection du format



#### Attention

Dans le domaine de la construction d'indicateurs en santé il existe un assez grand nombre de présentations qui matérialisent différentes méthodes adaptées à la solution du problème. La façon dont vont être formalisées les questions est extrêmement importante dans la mesure où ce format conditionne les étapes ultérieures de construction et de validation de la mesure.

Il est relativement complexe de faire un inventaire exhaustif des différents formats. Nous nous attarderons à ceux qui sont le plus souvent utilisés dans les constructions d'indicateurs de santé.

Nous distinguons classiquement deux grandes classes : les échelles à modalité catégorielle, les échelles de jugement continu.

#### *Echelles de type catégorielle*

Les échelles de type catégorielle sont de deux types :

### ◆ Echelle dichotomique (D'accord - Pas d'accord) :

Ces échelles consistent en de simples déclarations d'agrément suite à une liste de questions posées. Dans le domaine de la santé ce sont essentiellement les échelles constituées de réponses oui - non. Le scorage de ce type d'échelle est simple. Il convient d'édicter une règle cohérente : par exemple oui = 1 et non = 0. La sommation simple donnera un score élevé d'accord ou un score bas de désaccord.



#### Exemple

Nous pouvons en donner l'exemple suivant, à partir de l'échelle de vécu de l'anesthésie présentée dans la seconde partie. Nous aurions pu dans cette échelle formuler les questions sous la forme :

	Oui	Non
Après l'intervention étiez-vous gêné pour faire votre toilette ?		
Après l'intervention étiez-vous gêné pour aller aux toilettes ?		
Après l'intervention étiez-vous gêné pour prendre vos repas ?		

↑ TAB. 2

Si ces questions permettent d'apprécier la dimension de la gêne, alors en recodant 1 = non et 0 = oui, une note élevée après sommation aurait indiquée un niveau de gêne important. La réponse faite aux différentes questions pourrait être pondérée.

### ◆ Echelles à niveaux de réponses multiples ou Echelles à catégories spécifiques :

Très souvent les instruments de mesure de la qualité de vie liée à la santé ont recours à des échelles présentées sous forme d'échelles à catégories spécifiques (énoncés).

Il a été longuement reproché au format de réponse à catégories multiples de ne pas être adapté au mode de questionnement des scientifiques, qui en pratique chercheraient à savoir si le sujet prend une position contre son contraire (d'accord - pas d'accord). Il en tient là d'une confusion entre l'information fournie par la mesure et la décision. Si nous retenons le précédent argument comme vrai, ce dont nous pourrions discuter, il est clair qu'une décision peut être prise à partir d'une échelle à catégories multiples dès lors que nous posons un cut point rendant cette échelle à deux modalités, alors qu'une décision prise à partir d'une échelle oui - non ne permettra pas en retour d'apporter la finesse du jugement produite par les échelles à catégories multiples. Le second reproche formulé était l'addition de catégories ne faisant au final qu'ajouter du bruit (ou de l'erreur) dans le processus de mesure, dans la mesure où les sujets ne seraient pas capables d'avoir un jugement plus fin ou plus précis que le simple jugement binaire. Il a été depuis largement démontré et reconnu que si cette assertion pouvait se révéler exacte sous certaines circonstances, elle était cependant généralement tout à fait injustifiée et injustifiable.



#### Exemple

Le sujet n'a que cinq choix de réponses.

	Pas du tout	Un peu	Modérément	Beaucoup	Enormément
Après l'intervention étiez-vous gêné pour faire votre toilette ?					

↑ TAB. 3

Cette modalité de réponse est encore appelé échelle de Likert(5). Elle consiste à mesurer l'attitude d'un individu à partir d'une série d'énoncés, puis d'agréger les scores obtenus pour chaque énoncé afin d'obtenir un score d'attitude. Le nombre d'énoncés est classiquement de cinq ou 7. L'énoncé médian, le troisième, ne représentant pas systématiquement une position neutre. La construction d'échelle de Likert doit respecter certaines règles énoncées par Crocker et Algina, applicables de notre point de vue aux autres formats. Leurs recommandations sont présentées au point 2 de ce chapitre (Adaptation - Ecriture).

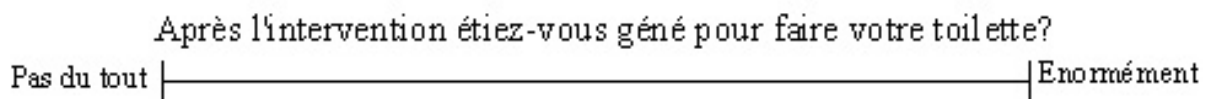
### *Echelles à jugement continu*

Les échelles à jugement continu s'imposent du fait même que nombre de variables d'intérêt, tout particulièrement dans le domaine des sciences biomédicales sont continues. Se pose alors la question de quantifier le jugement de celles-ci. Les échelles graphiques diffèrent des échelles à catégories spécifiques, dans le sens où le sujet a la liberté d'exprimer son attitude sur un continuum, classiquement appelé Échelle Visuelle Analogique (EVA) ou Échelle Analogique Visuelle (EAV).



### **Exemple**

Pour reprendre notre exemple nous aurions pu alors avoir pour le même énoncé :



↑ IMG. 1

Le sujet n'a alors qu'à se positionner sur le continuum qui sépare les deux extrêmes en général représentant 10 cm. Ensuite nous pourrions mesurer l'attitude du sujet, à l'aide d'une règle graduée.

L'utilisation de ces échelles est particulièrement appréciée des médecins qui voient en cette approche une démarche rapide, précise, dont les résultats sont fortement corrélés avec les appréciations faites par les échelles à catégories spécifiques.

Ces avantages sont plus perçus que réels. Le niveau de précision fourni, classiquement 1% sur une échelle 0 - 10 à une décimale, semble plus être une illusion de construction qu'un degré d'exactitude riche de sens. Quel est le niveau d'attribut sous jacent mesuré entre 2,3 et 2,4 ? La simplicité invoquée est aussi largement discutée par différents auteurs qui montrent que cette modalité de présentation peut poser des problèmes réels pour certains sujets (6). Ce mode de présentation est par exemple inadapté chez les personnes âgées, et nous savons que si nous souhaitons cependant y recourir il faudra présenter ces échelles non pas horizontalement mais verticalement. Enfin nous comprenons que cette façon d'apprécier l'attitude peut être relativement lourde dès lors que le questionnaire comporte une vingtaine de questions présentées sous ce format passé chez 100 sujets (soit 2000 mesures). Cet aspect constitue la principale limite à l'utilisation de ce type de format dans le développement d'indicateurs de QVLS. Au final ce type de présentation, malgré l'adhésion des praticiens, est très peu recommandée, de telle sorte que l'utilisation de ces échelles est donc relativement rare dans le domaine de la QVLS, afin de construire des indicateurs synthétiques.

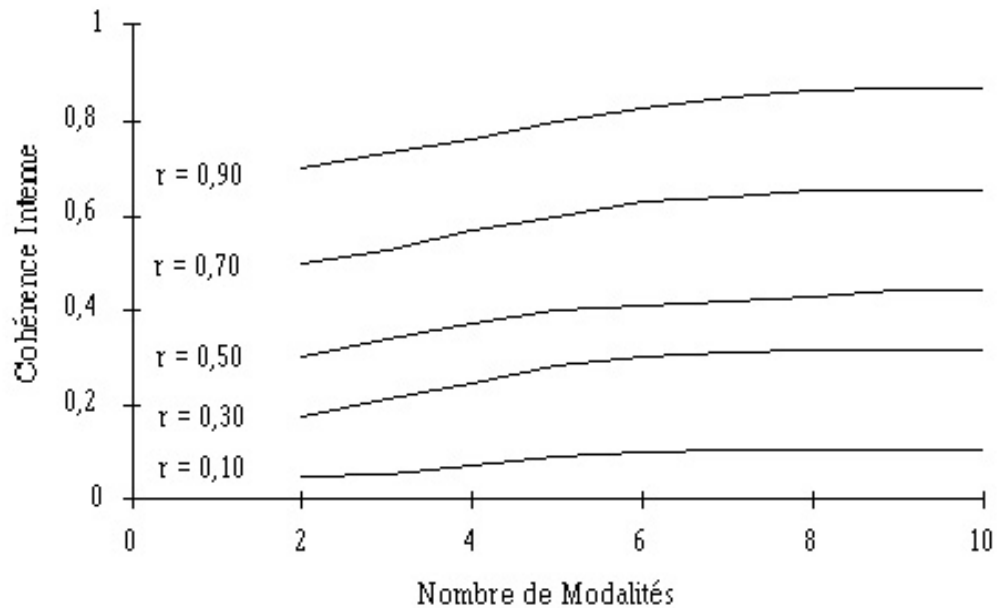


### Remarque

Le recours à ce type de format, nous amène à nous poser un certain nombre de questions afin de maximiser la précision de la mesure opérée et d'en diminuer les biais.

#### ◆ Nombre de modalités de réponses ?

Retenir un nombre de modalités inférieur à la capacité de discrimination du cotateur, qu'il soit médecin ou patient, implique une perte d'information. Cette capacité de discrimination est en fait très dépendante des situations dans lesquelles nous nous trouvons, et il n'est pas toujours évident de la présumer. Nishito et coll. (7) montrent au travers de simulations qu'ils mènent, que plus le nombre de modalités croît plus la cohérence interne entre deux variables augmente en tendant de façon asymptotique vers un plateau. Ce plateau est atteint au-delà de 5 à 7 modalités, selon le niveau de corrélation entre les variables comme l'illustre la figure suivante.



Ces résultats sont confirmés par de nombreux auteurs, qui tous recommandent de recourir à un nombre de catégories variant entre 5 et 7.

#### ◆ Nombre maximal de modalités de réponses ?

D'un point de vue purement statistique, nous avons vu que l'augmentation du nombre de modalités s'accompagne d'une augmentation asymptotique de la cohérence. Cependant en pratique il est évident qu'un nombre trop élevé de catégories affectera la cotation. Streiner recommande de ne pas dépasser 10 à 15 modalités (8) ce qui nous semble déjà trop important.

#### ◆ Nombre de modalités de réponses pair ou impair ?

Cette question est souvent débattue sans qu'aucune réponse univoque ne puisse être apportée (8). Le nombre de présentation impair offre la possibilité de proposer une position neutre, dès lors que celle-ci est intéressante à documenter, tout en sachant qu'il existe une tendance naturelle à se positionner à la moyenne et donc à voir la variabilité de la réponse diminuer. Inversement le mode de présentation pair, "forcera" l'individu à se positionner, mais perdra l'information parfois pertinente du sujet en situation d'indifférence quant à la formulation. Le choix doit être fonction des objectifs poursuivis (8), cependant pour notre part nous pensons qu'à moins d'avoir une solide argumentation en faveur de la modalité paire, la modalité impaire permettra de substituer les inévitables valeurs manquantes par la réponse centrale qui peut présenter des avantages lors du traitement de celles-ci.

#### ◆ Ordre ?

Certains auteurs se sont interrogés sur l'opportunité de générer de façon aléatoire l'ordre des modalités de réponses, de telle sorte que celles-ci n'aillent pas de façon systématique du plus mauvais au meilleur, ou du plus souhaitable au moins souhaitable. Le souci de ces auteurs est de minimiser les biais, notamment celui de la tendance à reproduire la réponse "yea saying "

que nous pouvons rapprocher à l'effet de halo. L'effet de halo est décrit à propos des questionnaires présentés en une colonne sur une seule page, où le sujet effectuera plus pour chaque formulation une cotation rapide sur la base d'une impression générale, que la cotation réelle de la question. Si ce souci peut paraître louable il semble très peu pratique il est en effet à craindre que les sujets interrogés ne remarquent pas par inattention, les changements de présentation, nous amenant à une collection de données ininterprétables. Pour notre part nous ne recommandons pas de recourir à la génération aléatoire de l'ordre.

## 2. Adaptation - Ecriture

Crocker et Algina (1) proposent de s'astreindre à respecter les règles suivantes, permettant de formuler les questions les plus adaptées possibles au sujet, quel que soit le contexte dans lequel sera développée la mesure.

- ◆ Mettre les énoncés au présent,
- ◆ Eviter les énoncés trop factuels,
- ◆ Eviter les énoncés pouvant avoir des interprétations variables,
- ◆ Eviter les énoncés insuffisamment discriminants qui donnent lieu à des réponses uni-modales de la part des sujets évalués,
- ◆ Essayer de mettre un nombre égal d'énoncés exprimant des sentiments positifs et négatifs,
- ◆ Faire des énoncés courts, n'excédant pas 20 mots,
- ◆ Mettre ces énoncés dans une forme grammaticale correcte,
- ◆ Eviter les énoncés comportant des terminologies trop universelles comme : tout, toujours, aucun, rarement,
- ◆ Eviter d'utiliser des qualificatifs indéfinis comme : seulement, juste, un peu
- ◆ Les énoncés doivent être exprimés simplement; éviter notamment les énoncés comportant les terminologies : si vous, à cause de ...,
- ◆ Utiliser un vocabulaire accessible aux personnes interrogées,
- ◆ Eviter les terminologies négatives : aucun, jamais ...

## 3. Revue des énoncés

### Recommandations



**Démarche**

La revue des énoncés consiste en la relecture de ceux-ci afin de corriger les erreurs susceptibles d'avoir été générées. Elle s'appuie sur les règles préalablement établies. Elle peut cependant s'opérer au regard d'une grille de lecture (1). Cette grille soulignera plus particulièrement les points suivants qui nous semblent pertinents :

- ◆ Vérifier la précision de l'énoncé,
- ◆ Vérifier le caractère approprié, la pertinence de la formulation au regard des objectifs poursuivis,
- ◆ Repérer les défauts de construction technique,
- ◆ Vérifier la construction grammaticale,
- ◆ Vérifier l'absence de biais dans la formulation, notamment par une formulation offensante ou embarrassante pour le sujet interrogé,
- ◆ Vérifier le degré de lisibilité.

Une attention toute particulière doit être portée quant à la référence temporelle. Il conviendra de choisir le temps de référence sur lequel nous interrogeons le sujet. Juniper et coll. rapportent qu'intuitivement, la référence aux deux semaines écoulées est une limite raisonnable, période durant laquelle le sujet se rappelle avec suffisamment de précision des différents sentiments ressentis (2). Cependant cette période de référence peut être modifiée en fonction des études, ou des arguments des experts cliniciens.

Une attention toute particulière doit également être portée quant à la lisibilité des énoncés, et ce tout particulièrement quant ces échelles seront administrées auprès d'enfants ou de jeunes adolescents. Il conviendra d'éviter les idiomes, le jargon, les métaphores posant des problèmes lors des adaptations dans d'autres langages et d'autres cultures. Juniper et coll. (2) rapportent différents exemples de formulations en anglais difficilement traduisibles et cependant rencontrées dans des questionnaires comme "down in the dumps" (avoir le cafard) ou "pooped" (crevé, vanné).

Par ailleurs certains auteurs rapportent que cette lisibilité devrait être évaluée de façon standardisée au moyen des tests de lisibilité classiquement utilisés dans d'autres domaines. Nous remarquerons que toutes les procédures de lisibilité classiquement rapportées ne sont pas adaptées à l'architecture générale des textes qui sont utilisés dans les échelles. En effet les intitulés sont généralement très courts.

## Réalisation

Cette revue peut se faire de différentes façons, sans que celles-ci soient nécessairement exclusives : relecture par des groupes d'experts, élaboration par des groupes de patients (focus group), par les familles des patients ... Le choix est surtout conditionné par les objectifs initiaux poursuivis. Il n'existe pas d'attitude préétablie idéale, ce choix devant être discuté et documenté dans un contexte donné.

Concernant cette revue d'items, nous pouvons nous interroger sur l'opportunité de sa réalisation avant le tri des items décrits au chapitre suivant. Cette question est de prime abord pertinente. Nous pourrions opérer cette revue, une fois le tri réalisé, rendant celle-ci moins fastidieuse, moins longue et donc moins coûteuse. En fait nous pouvons objecter que la sélection qui s'opérerait alors ne s'effectuerait pas nécessairement sur les énoncés les mieux formulés. Certains énoncés sont rejetés uniquement pour des raisons de formulation et auraient pu être corrigés si cette revue s'était inscrite plus précocement. Il s'agit là, de convenance personnelle. Il est certain que le concepteur disposant de moyens suffisants doit entreprendre cette revue tant avant qu'après le tri. La seule règle que nous puissions retenir est que, dès lors que nous suspectons que la validité de contenu peut poser des problèmes il est alors souhaitable de faire revoir les items sélectionnés après tri.

## Chapitre D. Tri des questions

### 1. Nombre de sujets

Avant d'entreprendre le passage de l'échelle de mesure auprès de larges échantillons en vue de sa validation, il convient de tester les énoncés sur des échantillons plus modestes de sujets. La question est de déterminer la taille de ces échantillons, qui au final est fonction de la taille de la population cible. Dès que la population cible est réduite 15 à 30 sujets seront admis comme suffisants. Inversement les tris peuvent être menés sur 200 à 300 sujets si la population cible est conséquente (1).

### 2. Réalisation

L'essentiel de cette étape est de repérer le comportement des sujets testés pour chacun des énoncés, et de recueillir tous commentaires utiles à la sélection ou à la reformulation de ces énoncés. Il est donc capital que l'équipe conceptrice de l'instrument soit en relation directe avec les sujets testés.



#### Attention

De notre point de vue, il existe au travers de cette démarche un biais difficilement contrôlable dès lors que l'échelle de mesure est prévue comme devant être auto administrée et remplie de façon isolée. La pertinence des informations recueillies est à pondérer par le fait que le test ne se déroule pas dans les conditions où il est prévu que celui-ci s'opère.

Cette période de test est suivie d'une période d'analyse par les concepteurs, ainsi que par toutes personnes jugées utiles comme définies au chapitre précédent : experts, patients, famille ... L'analyse sera aussi documentée de statistiques descriptives pour chaque énoncé, afin de vérifier que les réponses apportées comportent suffisamment de variabilité inter-sujets nécessaire à la construction.

Patrick propose de décomposer cette étape de tri préalable en deux étapes : une étape de prétests, et une étape d'étude pilote (3).

- ◆ Les prétests sont des études menées sur un nombre restreint de sujets, 5 à 25, afin d'estimer le temps de passage, les réactions des sujets, la compréhension des instructions et des questions, la présentation. Par ailleurs, il conviendra de recueillir l'opinion des sujets et des éventuels enquêteurs.
- ◆ Les études pilotes sont conduites sur un nombre plus conséquent de sujets (>25). Elles explorent les points préalablement cités. Elles permettent par ailleurs de tester certaines hypothèses avant la réalisation de l'étude, de comparer différentes méthodes d'administration ou de mesure, de tester la réalisation, de collecter des données pour l'étude des propriétés métrologiques.

Ces études peuvent être répétées (itérations successives) jusqu'à ce que nous n'identifiions plus de changement à opérer (2).

Ce n'est qu'après cette ultime étape que le questionnaire final est élaboré et expérimenté auprès de larges populations. Les données recueillies permettront alors d'explorer les propriétés statistiques du score qui seront développées dans les chapitres ultérieurs exposant notamment l'analyse d'items qui peut être utilement développée dès la phase de tri des questions.

L'identification des populations auprès desquelles le questionnaire devra être proposé ne pose pas de problèmes très spécifiques, dès lors que notre stratégie de développement respecte la rigueur des études développées en recherche clinique et/ou épidémiologiques. Cependant quatre caractéristiques essentielles doivent être isolées et explorées :

- ◆ Souhaite-t-on évaluer des sujets sains ou des sujets malades ?
- ◆ Souhaite-t-on évaluer des sujets jeunes ou des sujets âgés ?
- ◆ Souhaite-t-on explorer des personnes possédant toutes leurs capacités cognitives ou non ?
- ◆ Souhaite-t-on travailler dans des groupes culturels ou ethniques particuliers ?

### 3. Réalisation de l'enquête et analyse

Il n'y a pas de spécificité à la conduite même d'une enquête explorant la qualité de vie. Comme dans toutes études, il conviendra de maîtriser les sources d'erreur.

Fowler(9) en isole quatre principales ayant un impact fort dans les études de qualité de vie :

- ◆ les biais d'échantillonnage, où la population obtenue est différente de celle souhaitée pour laquelle la mesure a été conceptualisée,
- ◆ les biais de compréhension, dus aux questions mal comprises,
- ◆ les biais dans la conduite de l'entretien ou du recueil des données,
- ◆ les biais classiques de décodage, de saisie et de scorage.

La plupart de ces biais sont contournables dès lors que des procédures de standardisation ont été prévues et écrites. Ceci est plus particulièrement vrai dans la conduite des entretiens. Il existe différents modes d'administrations comme l'observation directe, l'entretien direct, l'entretien par téléphone, l'autoquestionnaire, l'approche par les proches. Ces méthodes ont des coûts et des impacts en termes de fatigue différents et doivent être discutées tant à la conception qu'avant toute utilisation.

La conduite d'entretien permet d'avoir des taux de réponses plus élevés que lors du recours à des autoquestionnaires notamment par voie postale. Par contre, ces derniers garantissent si besoin l'anonymat du sujet. Certains auteurs proposent de recourir aux autoquestionnaires sous surveillance d'un enquêteur afin de minimiser le taux de non réponse (3). Ainsi un concepteur doit s'interroger sur l'opportunité future d'utiliser l'auto-questionnaire mis au point lors d'entretien téléphonique.

Pour certains auteurs le contrôle de ces biais, présente peu d'intérêt dans la mesure où l'instrument développé a pour seul objectif de différencier les populations explorées, ces biais se répartissant de façon équitable dans les différents groupes alors considérés. Cette approche du problème est pour le moins empirique. L'objectif de la construction de la mesure demeure dans le contrôle des biais afin d'en maximiser la validité. A ce titre en dehors des biais classiquement rapportés par Fowler (9) nous pouvons mentionner : Biais de réponse socialement désirable; Biais de déviation; Biais de d'acquiescement et biais d'opposition; Biais d'aversion des extrêmes; Biais de distribution des réponses; Biais du à l'effet de halo ...

## Chapitre E. Validation, Acceptabilité et recommandation

### 1. Validation

La validation doit répondre à trois questions : l'instrument est-il pertinent ou valide, fiable et applicable ? Un outil ne peut être considéré comme fonctionnel que s'il présente des qualités de mesure suffisantes.

Nous exposerons les différentes composantes traditionnellement reconnues de la validité : validité de contenu, validité de construit, validité de critère. Nous verrons que les deux premières sont décrites de façon relativement univoque par les différents auteurs. La validité de construit, plus récemment identifiée, est à l'origine de multiples débats. Deux approches différentes du processus de validation coexistent et se recoupent. La première individualise quatre composantes : validité de contenu, validité sur critère, validité de structure (interne et externe) et validité de prédiction; la seconde scinde l'ensemble du processus de validation en deux : validation interne et validation externe.



### La validité

La validité (ou pertinence) d'un instrument est sa capacité à bien mesurer ce qu'il est censé mesurer.

La validité du contenu consiste à juger si le jeu de questions sélectionnées représente bien toutes les facettes du concept à mesurer. Il n'existe pas d'indice pour nous assurer de la validité de contenu. Il s'agit bien plus d'une démarche : s'assurer qu'en terme de contenu, les instruments développés soient représentatifs de ce que l'on tente de mesurer.

Cette notion de représentativité renforce l'idée d'un lien étroit entre validité du contenu et échantillonnage. Il est donc impératif, pour contrôler la validité du contenu d'un instrument : de spécifier le plus clairement possible l'univers duquel l'instrument est issu, d'adopter un processus d'échantillonnage rigoureux à l'intérieur de cet univers, d'identifier les composantes de l'instrument et finalement, d'être capable de porter un jugement objectif sur cet instrument développé (1).

*La validité de structure interne* explore la cohérence de l'agencement interne de l'échelle, notamment par l'analyse des corrélations des réponses aux questions entre elles ou par la mise en oeuvre de techniques d'analyses de données. En 1954, le terme de "construct validity" est introduit pour la première fois dans les recommandations de l'APA Commitee. En 1960, Campbell montre que toutes les mesures élaborées en psychologie ne peuvent être appréhendées par les classiques validité de contenu et validité de critère, alors que tout laisse à penser que ces mesures sont valides. En effet, bon nombre des tests développés notamment dans le domaine de la personnalité depuis 1920 et considérés comme valides, ne répondent pas au critère de validité pratique, en raison de l'absence de gold standard dans ce domaine. Pour ces tests, c'est donc la validité de construit qui doit être plus particulièrement étudiée.

Campbell (10) distingue deux types de validité de construit. Le premier, appelé validité de trait, est applicable pour des instruments dont par théorie on suppose seulement qu'ils décrivent un syndrome, un trait, ou une dimension de personnalité. Le second type, plus récemment isolé (11, 12), est appelé validité nomologique ou "nomological validity". Elle est en cause quand les concepteurs tentent de vérifier une (ou plusieurs) hypothèse(s) générée(s) à partir d'un cadre théorique formel au sein duquel est développé l'instrument à valider.

Un autre type de validation, la validité factorielle, est souvent utilisé par les concepteurs contemporains. Elle étudie la structure interne de l'instrument par des méthodes d'analyse factorielle, c'est-à-dire le regroupement des items d'un questionnaire selon différentes dimensions du concept mesuré. Elle pourrait par exemple constater que les réponses aux items d'un questionnaire de qualité de vie liée à la santé s'organisent en trois dimensions : physique, psychique, sociale.

La validité de structure interne porte sur la logique interne de l'instrument, sa logique de construction. En pratique l'exploration de cette propriété est basée sur :- l'analyse d'items; - l'étude des convergences item-échelle; - l'étude de la discrimination item-échelle; - l'étude des corrélations des différentes dimensions de l'échelle; - l'analyse factorielle.

*La validité de structure externe* repose soit sur l'étude des corrélations de cette mesure avec une autre échelle de QVLS faisant référence, soit sur l'étude des liaisons avec d'autres descripteurs externes : socio-démographiques, cliniques...

C'est un processus de validation de l'échelle dans lequel on a recours à une variable externe, un autre instrument de mesure, ou d'autres concepts .... On distingue selon qu'il existe ou non un critère posé en référence absolue (gold standard) la validité de structure externe référentielle de celle non référentielle.

- ◆ Validité de Structure Externe Référentielle : Selon que l'observation du critère est contemporaine de la mesure par l'échelle de qualité de vie, ou situé dans l'avenir, on parle alors de : Validité concourante ou concomitante et de Validité prédictive,
- ◆ Validité de Structure Externe non Référentielle : Elle peut être appréhendée par les différentes méthodes exposées et comporte les éléments suivants : Validité de Trait convergente ou discriminante; Validité Nomologique : tests d'hypothèses portant sur des variables externes ou des concepts indépendants.



### **La fidélité**

La fidélité (ou précision) de l'instrument est principalement appréciée sur deux critères : la cohérence interne appréciée par le calcul du coefficient alpha de Cronbach (des valeurs au-delà de 0,70-0,80 sont considérées comme souhaitables), la reproductibilité qui est sa capacité à produire des scores de QVLS comparables lorsque la mesure est répétée, alors que l'état de l'individu est stable. Quand l'instrument fait appel à un évaluateur externe, l'étude de la fiabilité doit comprendre l'étude de la fidélité intercotateurs (capacité à produire des scores similaires quand à un temps donné au moins deux évaluateurs différents cotent le même sujet avec le même instrument).



### **La sensibilité au changement**

La sensibilité au changement est la capacité de l'instrument à mettre en évidence une variation jugée comme pertinente par les experts. Ce dernier critère est particulièrement requis dans un contexte d'évaluation d'interventions de santé, comme lors d'un essai thérapeutique.

## 2. Acceptabilité

Un instrument sera applicable si les conditions d'utilisation montrent que le coût de mise en oeuvre est modeste, que l'acceptabilité par les sujets et la communauté médicale est grande, et que le temps de passation est faible.

L'acceptabilité d'un instrument peut être appréciée de différents points de vue:

- ◆ Nous pouvons nous interroger sur le fait qu'il réponde aux objectifs de la personne le mettant en oeuvre et de ses ressources. Auquel cas il conviendra de développer et de sélectionner l'instrument sur le critère de sa longueur, dans la mesure où celle-ci peut perturber les autres évaluations à développer. Le coût sera aussi un critère pertinent, dû au nombre d'enquêteurs, au temps en plus, à la complexité du recueil, à l'analyse des données. L'applicabilité éprouvée dans la même situation sera enfin prépondérante.
- ◆ Le second point de vue à documenter est celui du sujet évalué, tout particulièrement le temps nécessaire et l'éventuel niveau de stress occasionné. Le temps nécessaire peut varier bien entendu en fonction de la longueur et de la complexité du questionnaire mais aussi en fonction de caractéristiques propres aux sujets comme l'âge, l'éducation, l'altération des fonctions cognitives susceptibles d'augmenter ce temps (13).
- ◆ Le troisième point de vue est celui de l'impact même du questionnaire sur l'enquêteur. Certains questionnaires nécessitant le recours à un enquêteur, impliquent une formation et un entraînement à la cotation afin d'augmenter la fiabilité des données. C'est le cas par exemple du Quality of Well Being Scale ou du SIP où les auteurs recommandent fortement cet entraînement (14).

## 3. Ecriture des recommandations d'utilisation pratique

Tout concepteur d'indicateurs de qualité de vie liée à la santé devra mettre à la disposition des futurs utilisateurs des recommandations d'utilisations. Ces recommandations préciseront les précédents points évoqués. Plus particulièrement devront être signalés les modalités d'administration, le mode de présentation des questions, la référence temporelle utilisée, l'acceptabilité et le coût, ainsi que les modalités de scorage. En effet le choix d'un outil consistant en une mesure composite nécessite que l'utilisateur potentiel ait accès aux algorithmes de scorage.

## Bibliographie

1. Crocker L, Algina J. Introduction to classical and modern test theory. New York. Holt, Rinehart and Winston ; 1986.

2. Juniper EF, Guyatt GH, Jaeschke R. How to develop and validate a new health related Quality of life instrument. In *Quality of life and pharmacoeconomics in clinical Trials*. Second Ed by B. Spilker. Lippincot-Raven Publishers ; 1996.
3. Patrick DL, Erickson P. Health status and Health Policy, Quality of life in health care evaluation and ressource allocation. Oxford University, New York, 1993.
4. Kirschner B, Guyatt G. A methodeological framework for assessing health indices. *J Chronic Dis* 1985 ; 38 : 27-36.
5. Likert R. A technique for the measurement of attitudes. *Arch of Psycholog* 1932 ; 140 : 44-53.
6. Huskisson EC. A technique for the development of attitude scales. *Educational and Psychological Measurement* 1952 ; 12 : 313-15.
7. Nishisato N, Rorii Y. Effects of categorizing continuous normal distributions on the product-moment corelation. *Japanese Psychological research* 1970 ; 13 : 45-9.
8. Streiner DL, Norman GR. *Health Measurement Scales*. New York : Oxford University Press ; 1989.
9. Fowler FJ, mangione TW. *Standardized survey interviewing*. Newbury Park. CA : Sage Publications ; 1990.
10. Campbell DT, Fiske DW. Convergent and discriminant validation by multitrait multimethod matrix. *Psychol Bull* 1959 ; 56 : 81-105.
11. Cronbach LJ, and Meehl PE. Construct validity in psychological tests. *Psychological Bulletin* 1955 ; 52 : 281-302.
12. Jessor R and Hammond KR. Construct validity and the Taylor Anxiety scale. *Psychol Bull* 1957 ; 54 : 161-170.
13. Bush JW. general health policy Model/Quality of Well Being (QWB) scale. In *Assessment of quality of life in clinical trials of cardiovascular therapies*. Wenger NK, Mattson ME, Furberg CD, Elinson J eds, New York : LeJacq ; 1984a : 189-99.
14. Bergner L, Hallstrom AP, Eisenberg M. Health status of survivors of out-of-hospital cardiac arrest six months later. *Am J Public Health* 1984 ; 74 : 508-10.

# Adaptation trans-culturelle d'instruments existants

## Préambule

Durant les 20/30 dernières années de nombreux questionnaires de qualité de vie (QV) ont vu le jour, utilisés avec un intérêt croissant dans le cadre d'études internationales notamment dans des essais cliniques. Ainsi, s'est imposée la nécessité de disposer de mesures de QV partageables au niveau international et autorisant les comparaisons entre pays mais aussi l'agrégation des informations à une échelle internationale.

Cependant, l'évaluation de la QV n'est pas indépendante du contexte culturel. Celui-ci va conditionner la façon dont les individus appréhendent leur vie, la valeur qu'ils accordent aux choses et leurs aspirations. La question de l'adaptation et de validité trans-culturelle se pose d'ailleurs non seulement lorsqu'il s'agit de partager une mesure de QV entre pays, mais aussi au sein d'un même pays, par exemple avec les populations migrantes (1). Si globalement, les membres d'une même société partagent la même culture, on peut identifier au sein de nos sociétés des groupes se distinguant du reste de la société du fait par exemple de caractéristiques communes socio-démographiques, ethniques ou religieuses, et qui vont par là même partager des valeurs et des aspirations se démarquant pour partie de celles plus largement répandues dans la société.

L'adaptation et la validation trans-culturelle des instruments de mesure de QV constituent donc un enjeu central dans le champ de la mesure de QV. Dans ce chapitre, seuls les aspects relatifs à l'adaptation trans-culturelle entre pays sont évoqués.

De façon simplificatrice, lors du développement de mesures de QV à l'échelle internationale deux types de situation sont possibles :

◆ **l'approche séquentielle :**

un instrument initialement développé dans une langue source est traduit secondairement dans d'autres langues, appelées langues cibles (2),

◆ **le co-développement :**

un questionnaire est développé de novo de façon " simultanée " ou " parallèle " dans plusieurs pays (3, 4).

Le processus de validation trans-culturelle permet d'obtenir une mesure utilisable pour des investigations sur la QV dans des études internationales mais aussi dans des contextes culturels différents au sein d'un même pays (17). Il autorise sous certaines contraintes des comparaisons inter-culturelles. Il permet d'inclure les groupes présentant des spécificités culturelles même s'ils sont minoritaires comme les populations immigrées, évitant ainsi le biais de ne représenter que la culture dominante d'un pays.

## Chapitre A. Concept d'équivalence

L'équivalence conceptuelle entre les questionnaires développés dans les différentes cultures est un pré-requis pour assurer la pertinence inter-culturelle. Mais au-delà de l'équivalence conceptuelle, trois autres dimensions d'équivalence peuvent être évaluées : l'équivalence opérationnelle, l'équivalence de construit ou d'échelle, l'équivalence métrique (5, 6).

### 1. Equivalence conceptuelle



**Attention**

La correspondance entre les mots (équivalence sémantique) est difficilement réalisable d'une culture à l'autre compte-tenu du vocabulaire et de la grammaire propre à chaque langue. Certaines expressions/idiomes traduites littéralement n'ont pas de sens dans une autre culture et des expressions propres à la culture ciblée conservant le sens des expressions doivent être trouvées (équivalence des expressions/idiomes). Certaines situations qui sont évoquées dans la culture de l'instrument d'origine peuvent ne pas être transposables à une autre culture, elles devront être remplacées par d'autres situations appropriées à cette culture tout en préservant l'objectif et le sens de la question (équivalence expérientielle). Enfin, le même exercice doit être appliqué pour certains concepts qui traduits littéralement ne permettent pas la même représentation d'une culture à l'autre (équivalence conceptuelle).

## 2. Equivalence opérationnelle

Les différents modes d'administration possibles pour le questionnaire sont comparés en termes de propriétés psychométriques.

## 3. Equivalence de construit ou d'échelle

Les propriétés psychométriques des différentes versions source et cibles sont comparées.

## 4. Equivalence métrique

Il s'agit d'étudier dans quelles mesure des individus présentant un même niveau de QV vont obtenir des scores de même niveau sur le continuum des scores de QV pour les différentes versions de l'instrument.

# Chapitre B. Les différentes approches

L'adaptation trans-culturelle d'un instrument d'une langue source ou véhiculaire dans une ou plusieurs langues cibles va ainsi comprendre deux grandes étapes : la traduction et l'évaluation psychométrique de l'instrument (validité, fiabilité, sensibilité au changement).

Le langage étant un élément central de la culture, la traduction d'un questionnaire d'une langue source dans une autre langue va constituer un élément central du processus d'adaptation transculturelle.

L'analyse des propriétés métrologiques se fait suivant les mêmes modalités que celles décrites précédemment. La cohérence interne de l'instrument, sa stabilité, ainsi que sa sensibilité au changement sont étudiées. Plus spécifiquement, la comparabilité du construit dans les différents pays peut être explorée par des analyses factorielles confirmatoires, vérifiant que les dimensions retrouvées sont bien celles décrites initialement, et constituées des mêmes items.

L'étude du Fonctionnement Différentiel de l'Item (DIF), encore peu répandue dans le champ de la mesure de QV, peut par ailleurs permettre d'évaluer l'équivalence entre les traductions. Elle explore si un item de la version source " fonctionne de la même manière " que ceux traduits dans les différentes langues cibles : les sujets ayant le même niveau de QV ont-ils la même probabilité de donner la même réponse quelle que soit la traduction utilisée (7).

Il conviendra aussi de vérifier si le scorage et notamment le système de pondération est adapté au contexte cible. Il pourra être nécessaire d'adapter les pondérations initiales au contexte culturel.

Quatre démarches d'adaptation d'un instrument vers une culture cible sont décrites dans la littérature (8).

## 1. Approche ethnocentrique

*L'approche ethnocentrique* consiste à réaliser la traduction de l'instrument vers la culture cible en supposant a priori que l'instrument est approprié pour celle-ci. Elle implique que :

- ◆ les dimensions du concept étudié par l'instrument source soient transférables dans la culture cible,
- ◆ les langages source et cible soient suffisamment proches pour permettre une traduction qui ait une signification,
- ◆ le mode de questionnement par l'instrument source soit applicable à la culture cible.



### Attention

Le risque de cette approche est d'avoir un instrument qui ne soit pas valide dans la culture cible.

## 2. Approche pragmatique

*L'approche pragmatique (dite "etic")* implique de rechercher des champs équivalents dans les deux cultures et de ne traduire que les questions explorant des concepts communs aux deux cultures.

## 3. Approche emic plus etic

*L'approche "emic plus etic"* consiste, dans un premier temps, à traduire les concepts de base équivalents dans les deux cultures ("etic"). Dans un second temps, les aspects spécifiques au concept de QVLS dans la culture cible ("emic") sont individualisés et des questions spécifiques sont développés pour les explorer. Ces aspects "emic" ne pourront donc être évalués que dans la culture cible.

## 4. Approche analogique

*L'approche analogique* peut être mise en oeuvre quand les deux cultures sont si différentes qu'il n'est plus possible de traduire directement l'instrument source. Il est alors envisageable d'interpréter les questions de l'instrument d'origine pour le recréer dans le langage cible, suivant la démarche de construction de l'instrument original. Le risque de cette approche est qu'au final la différence entre les deux instruments pourra être telle, qu'en fait ils ne mesurent pas le même concept.

## Chapitre C. Déroulement du processus de traduction

Le processus de traduction, largement décrit dans la littérature, peut être résumé schématiquement en 4 étapes (9). Chaque étape participe à la qualité de la version finale en termes d'équivalence des concepts explorés entre la version source et la version finale (cible). Le processus de traduction doit concerner les questions mais aussi les modalités de réponses et les instructions assujetties au questionnaire. Il est important d'y associer les auteurs du questionnaire, qui pourront notamment aider à clarifier les concepts mesurés et donc améliorer la qualité de la version finale. Leur participation permettra également d'éviter la multiplication des traductions d'un même questionnaire dans une même langue, source de confusion pour les utilisateurs.

### 1. Traduction et retro-traduction

Plusieurs traductions de la langue source vers la langue cible sont réalisées indépendamment par des traducteurs dont la langue maternelle est la langue cible et ayant une pratique quotidienne de celle-ci.

Deux écoles s'affrontent quant aux recommandations à leur transmettre :

- ◆ l'une préconise d'expliquer aux traducteurs les objectifs et concepts explorés par l'instrument à traduire pour avoir une restitution plus fiable de ce que l'instrument est censé mesurer (10),
- ◆ l'autre recommande de ne fournir aucune explication afin de mettre à jour les éventuelles significations inattendues de l'instrument original (11, 12).



#### Démarche

- ◆ Les différentes traductions produites sont ensuite confrontées par un comité réunissant toutes les personnes dont le jugement permettra d'éclairer la traduction. Il s'agit des experts dans la pathologie et le concept à étudier, ainsi que des méthodologistes, ayant une connaissance suffisante des deux langues.
- ◆ Chaque item traduit est évalué en termes de contenu, signification, clarté des expressions, comparabilité avec l'échelle source, divergences des traductions. Toutes les objections sont discutées, et les items sont éventuellement réécrits, afin d'établir une traduction consensuelle dans la langue cible.
- ◆ La qualité de la traduction obtenue peut être évaluée en recourant à quelques (au moins deux) personnes dont la langue maternelle est la langue cible n'ayant pas participé aux étapes précédentes de traduction. Pour chaque item, ils jugeront de l'équivalence conceptuelle avec la version originale, la clarté et la pertinence du registre de langage utilisé. Les items ne satisfaisant pas à ces critères doivent être retraduits et réévalués.



### La rétro-traduction

La rétro-traduction est la traduction de la langue cible vers la langue source.

Elle peut être réalisée sur la version harmonisée ou sur toutes les versions traduites, par un traducteur dont la langue maternelle est la langue source et qui ne connaît pas la version source. La version rétro-traduite est comparée à la version source et, comme dans l'étape précédente, les divergences sont discutées par un comité constitué de plusieurs personnes issues d'horizons différents.

## 2. Etablissement du questionnaire final



### Démarche

L'élaboration de la version finale est basée sur les différentes traductions et contre-traductions et sur leur analyse. Des modifications peuvent être réalisées dans la version cible pour en améliorer l'équivalence avec la version source, il faut alors à nouveau procéder à une rétro-traduction. Les instructions ou la présentation de l'instrument pourront être modifiées, des items inappropriés ou ambigus modifiés ou rejetés, et de nouveaux items générés si cela est nécessaire. De même, les modalités des réponses de chaque question pourront être revues, la traduction devant par exemple maintenir l'ordinalité ou les propriétés d'intervalle des échelles de réponses. Pour résoudre les divergences le comité utilisera les données qu'il a collectées. Il pourra décider d'entreprendre à nouveau le processus traduction contre-traduction, jusqu'à ce que la version finale soit acceptée et validée. Ce processus itératif peut être en pratique lourd à mettre en oeuvre.



### Démarche

Une autre méthode a été proposée pour améliorer l'adaptation trans-culturelle, consistant à considérer la version source et la version finale comme aussi importantes (13). C'est notamment le cas du co-développement simultané entre pays où la version du questionnaire dans la langue source ne doit pas avoir un statut différent des autres versions en dehors du fait qu'elle constitue celle de la langue véhiculaire, généralement l'anglais. Les deux versions sont sujettes à des modifications éventuelles pendant la procédure de traduction. Cette méthode doit être de préférence conduite en relation étroite avec les concepteurs de l'échelle, qui pourront, si des items posent des problèmes, fournir une version de travail de l'échelle ou de certains items pour maintenir le contexte de la question. Rechercher une manière commune d'exprimer un concept dans les deux langues est la meilleure façon de maintenir la validité de contenu dans la version finale.

## 3. Pré-test



### Pré-test

C'est l'évaluation de la clarté, la compréhension et l'acceptabilité du questionnaire. Elle doit être menée auprès de personnes, dont le profil est le plus proche possible des populations cibles.

Deux techniques sont proposées :

- ◆ Pour chaque item, les individus interrogés vont préciser, dans une question ouverte, le sens de leur réponse (14). Cette méthode permet de s'assurer que chaque item de la version finale est compris de manière équivalente à l'item source.
- ◆ Un groupe de personnes bilingues répond à la version source et à la version finale du questionnaire afin d'évaluer, pour chaque item, l'équivalence des réponses obtenues entre les deux versions. Cette technique est généralement difficile à mettre en oeuvre, mais peut être utilisée assez facilement dans les populations immigrées (15, 16).

Chaque question sera ensuite reprise une à une, commentée et analysée par le comité de pilotage. Si besoin, les remarques des participants seront intégrées dans le questionnaire après accord du comité de pilotage et de l'auteur. La version obtenue est appelée version finale, c'est celle qui sera soumise à l'exploration des propriétés métrologiques.

## 4. Harmonisation internationale

Lors du développement d'un questionnaire dans plusieurs langues cibles, notamment lors du co-développement d'un nouvel instrument entre plusieurs pays, une étape supplémentaire d'harmonisation internationale est nécessaire. Elle a pour objet de comparer l'ensemble des versions traduites et la version originale. Cette étape doit impliquer pour chaque langue cible des personnes dont la langue maternelle est l'une des langues cibles et bilingue dans la langue source. Cette harmonisation révèle les divergences éventuelles entre expressions ou concepts dans les différents pays.

Un rapport de traduction doit être établi, détaillant le processus de validation linguistique : la méthodologie employée, les difficultés rencontrées (traduction, concepts...), les solutions proposées et adoptées. Les différentes versions du questionnaire sont incluses dans le rapport.

## Chapitre D. Bibliographie

### 1. Références, citées dans le texte : lecture recommandée

6. 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(120) : 1417-32.

7. Qol 2003 Petersen MA, Groenvold M, Bjorner JB and al. Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Qual Life Res* 2003 ; 12 : 373-85.
8. Sartorius N, Kuyken W. Translation of Health Status Instruments. In : Orley J, Kuyken W, Eds. *Quality of Life Assessment : International Perspectives*. Springer-Verlag Berlin Heidelberg 1994.
9. Acquadro C, Jambon B, Ellis D, Marquis P. Language and translation issues. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
17. Bullinger M, Anderson D, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments for minimal requirements to optimal models. *Qual Life Res* 1993 ; 2 :451-59.

## 2. Articles cités dans le texte

1. Guarnaccia PJ. Anthropological perspectives : the importance of culture in the assessment of quality of life. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
2. Aaronson NK, Acquadro C, Alonso J, Apolone G, et al. International quality of life assessment (IQLA) project. *Qual Life Res* 1992 ; 1 :349-51.
3. Aaronson NK, Cull A, Kaasa S, Sprangers MAG. The EORTC modular approach to quality of life assessment in oncology. *Int J Ment Health* 1994 ; 23(2) : 75-96.
4. Sartorius N. A WHO method for the assessment of health related quality of life (WHOQOL). In : Walker SR, Rosser RM, eds. *Quality of life assessment : key issues in the 1990's*. Boston : Kluwer Academic, 1993 ; 201-7.
5. Hui C, Triandis HC. Measurement in cross-cultural psychology : a review and comparison of strategies. *Cross Cultural Psychol* 1985 ;16 : 131-52.
10. Patrick DL, Erickson P. *Health Status and Health Policy. Quality of life in health care evaluation and ressource allocation*. Oxford University Press 1993.
11. American Psychological Association, American Educational Research Association and National Council on Measurement Used in Education. *Standards for educational and psychological tests and manuals*. Washington, DC American Psychological Association, 1966.
12. Guillemin F, Briançon S, Pourel J. Validity and discriminant ability of a French version of the Health Assessment Questionnaire in early RA. *Disabil Rehabil* 1992 ; 14 : 71-77.
13. Werner O, Campbell DT. Translating, working through interpreters, and the problem of decentering. In : Narroll R, Cohen E, Eds. *A Handbook of Method in Cultural Anthropology*. New York ; American Museum of Natural . History Press : 1970 : 398-420.

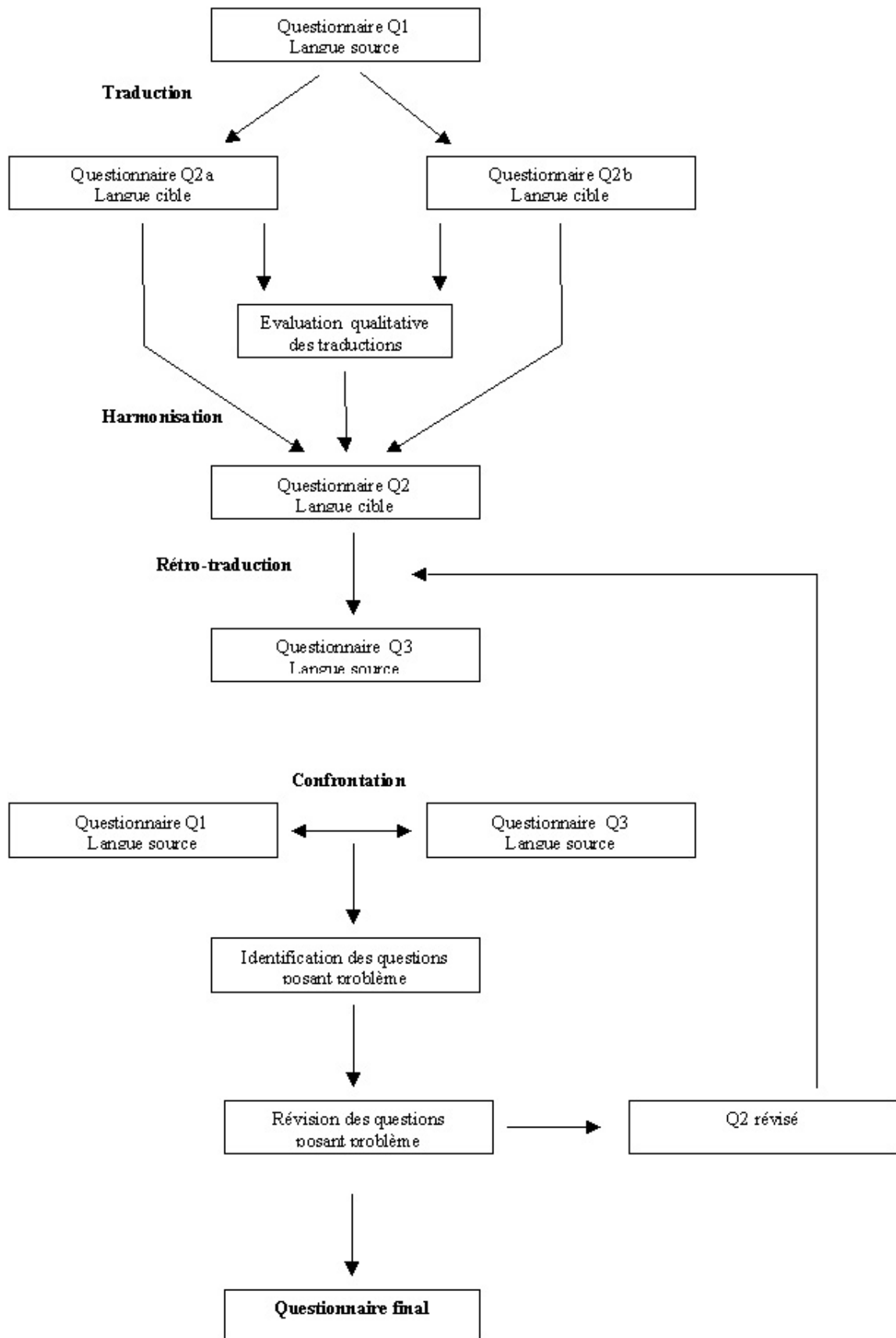
14. Schuman H. The random probe : A technique for evaluating the quality of closed questions. *Am Sociol Rev* 1966 ; 31 : 218-222.
15. Brislin R, Lonner W, Thorndike R. *Cross cultural research methods*. New York : Wiley ; 1973 : 32-81.
16. Spielberger CD, Sharma S. Cross-cultural measurement of anxiety. In Spielberger CD, Diaz-Guerrero R, Eds, *Cross-cultural Anxiety*. Washington ; Hemisphere Publ Corp 1976 : 13-25.

### 3. Non cités dans le texte : pour plus d'information

- ◆ Power M, Harper A, Bullinger M. The World Health Organization WHO-QOL-100 : tests of universality of quality of life in 15 different cultural groups worldwide. *Health Psychol* 1999; 18: 495-505.
- ◆ Sprangers MA, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer Approach to quality of life assessment : guidelines for developing questionnaire modules. EORTC Study Group on Quality of Life. *Qual Life Res* 1993; 2(4): 287-95.
- ◆ Schmidt S, Bullinger M. Current issues in cross-cultural quality of life instrument development. *Arch Phys Med Rehabil* 2003 ;84( S29-S33).
- ◆ Anderson RT, McFarlane M, Naughton MJ, Shumaker SA. Conceptual issues and considerations in cross-cultural validation of generic health-related quality of life instruments. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
- ◆ Johnson TM. Cultural considerations. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
- ◆ Hunt S, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S. Cross-cultural adaptation of health-related quality of life measures. *Health Policy* 1991 ;19 :33-44.
- ◆ Anderson RT, Aaronson NK, Wilkin D. Critical review of the international assessment of health-related quality of life. *Qual Life Res* 1993 ;2 :369-95.
- ◆ Skevington SM, Tucker C. Designing response scales for cross-cultural use in health care : data from the development of the UK WHOQOL. *Br J Med Psychol* 1999 ;72 :51-61.
- ◆ Bullinger M. The challenge of cross-cultural quality of life assessment. *Psychol Health* 1997 ;12 :815-25.
- ◆ Custers JW, Hoijsink H, van der Net J, Helders PJ. Cultural differences in functional status measurement : analyses of person fit according to the Rasch model. *Qual Life Res* 2000 ;9 :571-8.

- ◆ Aaronson NK, Acquadro C, Alonso J, Apolone G, et al. International quality of life assessment (IQOLA) project. *Qual Life Res* 1992 ; 1 : 349-51.
- ◆ Aaronson NK, Cull A, Kaasa S, Sprangers MAG. The EORTC modular approach to quality of life assessment in oncology. *Int J Ment Health* 1994 ;23(2) :75-96.
- ◆ Sartorius N. A WHO method for the assessment of health-related quality of life (WHOQOL). In : Walker SR, Rosser RM, eds. *Quality of life assessment : key issues in the 1990's*. Boston : Kluwer Academic, 1993 ;201-7.
- ◆ Leplège A. Adaptation transculturelle d'instruments existants. In : Leplège A, Coste J. *Mesure de la Santé perceptuelle et de la qualité de vie : méthodes et applications*. Paris : ESTEM, 1999.
- ◆ Hutchinson JF. Quality of life in ethnic groups. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
- ◆ Anderson RT, Aaronson NK, Leplège A, Wilkin D. International use and application of generic health-related quality of life instruments. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.
- ◆ Cella DF, Lloyd SR, Wright BD. Cross-cultural instrument equating current research and future directions. In : *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Edition, B.Spilker eds. Philadelphia : Lippincott-Raven Publishers, 1996.

#### 4. Schéma : processus de traduction



↑ IMG. 3