

## ***La qualité de vie relative à la santé***

Cette partie présente la définition de la qualité de vie et ses particularités, la spécificité de la qualité de vie relative à la santé, les domaines nécessairement pris en compte pour l'évaluation de la qualité de vie et enfin l'importance du modèle conceptuel de Wilson et Cleary [2] pour comprendre les liaisons entre les différents domaines et par conséquent les éléments importants à prendre en compte.

Notez qu'à partir du paragraphe 3 du présent chapitre, nous n'utiliserons plus que la formulation qualité de vie pour parler de qualité de vie relative à la santé.

### **Définition de la qualité de vie**

C'est à partir de la définition de la santé donnée par l'Organisation Mondiale de la Santé (OMS) en 1948, définissant la santé comme étant « un état complet de bien-être physique, mental et social, et qui ne consiste pas seulement en l'absence de maladie ou d'infirmité » [3] que l'intérêt pour l'évaluation de la qualité de vie des patients a pris de l'ampleur. En effet, l'OMS replace la santé dans un contexte plus général. Jusqu'alors, la santé n'avait été étudiée que sous des aspects objectifs, cliniques ou biologiques (symptômes, événements indésirables, résultats biologiques...). Elle explorera désormais des aspects plus subjectifs (le ressenti vis-à-vis du traitement, la vision de soi...). Ainsi, la santé n'est plus seulement liée à la maladie ou à l'invalidité mais également au bien-être physique, mental et social. Les cliniciens ne doivent plus travailler dans l'unicité de la maladie mais dans la pluridimensionnalité du patient.

À partir de cette nouvelle approche de la santé, l'évaluation de la qualité de vie prend tout son sens. De nombreuses équipes de chercheurs vont travailler sur le sujet et présenter leur propre définition de la qualité de vie. Il n'existe donc pas de définition universelle précise de la qualité de vie.

Selon les équipes, la qualité de vie est définie comme :

- « Une satisfaction individuelle globale avec la vie et le sens général de son bien-être personnel »[4] ;
- « La perception subjective de satisfaction ou bonheur avec la vie dans les domaines importants pour l'individu » [5] ;

- « L'évaluation de satisfaction que font les patients de leur niveau actuel de fonctionnement par rapport à ce qu'ils perçoivent comme possible ou idéal »[6] ;
- « Un large éventail d'expériences humaines liées au bien-être de chacun. [...]La qualité de vie est définie par des expériences, des états et des perceptions subjectives [...]. Le terme « qualité de vie » va au-delà de la santé à proprement parler, mais elle est influencée par la perception que l'individu a de son environnement physique, psychologique, social, économique et politique » [7].

Ces définitions ont cependant toutes en commun d'intégrer des considérations objectives et des considérations subjectives. De ces variantes se dégage un cadre consensuel conformément au cadre global défini en 1948 par l'OMS. Il n'est guère étonnant alors que la définition la plus souvent évoquée soit celle proposée par l'OMS en 1993 : « la perception qu'un individu a de sa place dans la vie, dans le contexte de la culture et du système de valeurs dans lequel il vit, en relation avec ses objectifs, ses attentes, ses normes et ses inquiétudes. C'est un concept très large qui peut être influencé de manière complexe par la santé physique du sujet, son état psychologique et son niveau d'indépendance, ses relations sociales et sa relation aux éléments essentiels de son environnement » [8].

Finalement, quelle que soit la définition que l'on retienne de la qualité de vie, la subjectivité de son appréciation est recherchée, dans la mesure où c'est le patient lui-même qui est le premier interprète de son ressenti, selon ses émotions et son environnement.

### **Spécificité de la qualité de vie relative à la santé**

Lorsque l'on parle de qualité de vie relative à la santé il s'agit de spécifier qu'il est question de la qualité de vie étudiée dans un contexte médical. C'est « l'impact de la santé, incluant la maladie et le traitement sur le bien-être social, physique et psychologique » du patient [9].

La qualité de vie relative à la santé est également définie comme étant « un état de bien-être qui repose sur deux composants : (1) la capacité à réaliser des activités journalières qui reflètent le bien-être physique, psychologique et social ; (2) la satisfaction du patient avec son niveau de fonctionnement, le contrôle de sa maladie et des symptômes liés à son traitement » [10, 11].

Ainsi le qualificatif relatif à la santé, qui peut sembler réduire le champ d'application de la qualité de vie, l'accroît au contraire, en insistant sur l'importance qu'il y a à considérer l'individu dans sa globalité contextuelle, dont la pathologie tout comme la médicalisation en sont de possibles composantes.

Dans la suite de ce travail, par commodité de langage, nous parlerons elliptiquement de qualité de vie pour évoquer la qualité de vie relative à la santé [12].

### **La qualité de vie : un concept multidimensionnel**

Si les auteurs diffèrent sur la définition précise de la qualité de vie, un large consensus est établi pour la définir comme un concept multidimensionnel (cf. § 1 et 2 précédents) [2, 7, 10]. Néanmoins, des divergences existent sur la structure qu'elle doit prendre. En effet, pour certains auteurs, l'évaluation de la qualité de vie se structure au minimum autour de trois dimensions, que sont les domaines du physique, du psychologique et du social [9, 10]. Pour d'autres équipes une évaluation de la qualité de vie doit impérativement reposer sur un domaine supplémentaire, la symptomatologie, permettant ainsi une évaluation de la qualité de vie selon quatre dimensions au minimum [13, 14].

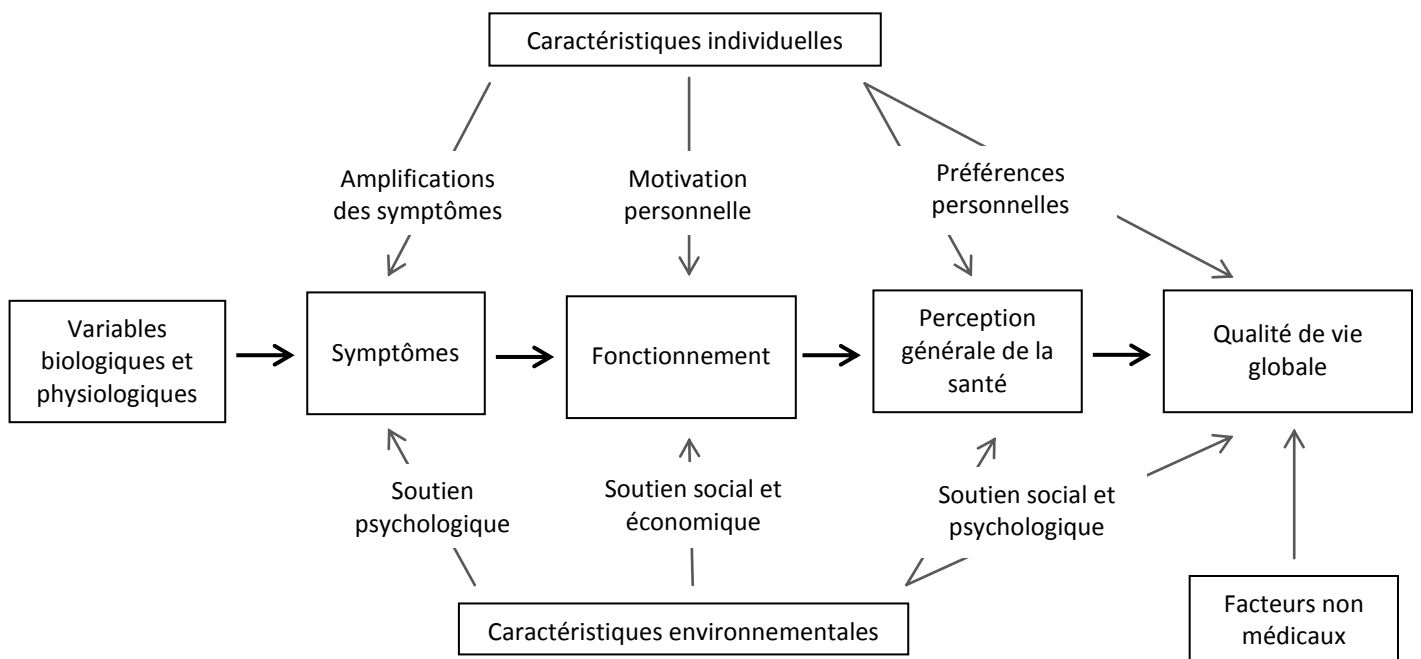
Ces propositions permettent de constituer les éléments fondamentaux à prendre en considérations pour évaluer la qualité de vie. Toutefois, il n'est pas interdit, il est même souhaitable, d'asseoir ces évaluations sur des domaines supplémentaires nécessairement sources d'enrichissement selon le sujet de recherche et conséquemment selon les axes d'investigations et d'analyse plus spécifiques choisis. Ainsi, l'image du corps, la fonction cognitive, la sexualité, les rapports avec l'équipe médicale, la politique, l'économie sont des domaines contribuant potentiellement à l'évaluation de la qualité de vie [2, 9, 15].

### **4. Modèles conceptuels de la qualité de vie relative à la santé**

Afin de mieux comprendre et visualiser les relations existantes entre les éléments pris en compte pour l'évaluation de la qualité de vie (domaines sus-cités, éléments environnementaux...), Wilson et Cleary ont développé un modèle conceptuel permettant d'appréhender les relations causales et les interactions existantes entre ces différents éléments et, ainsi, facilitant la méthodologie des interventions cliniques relatives à la qualité de vie [2].

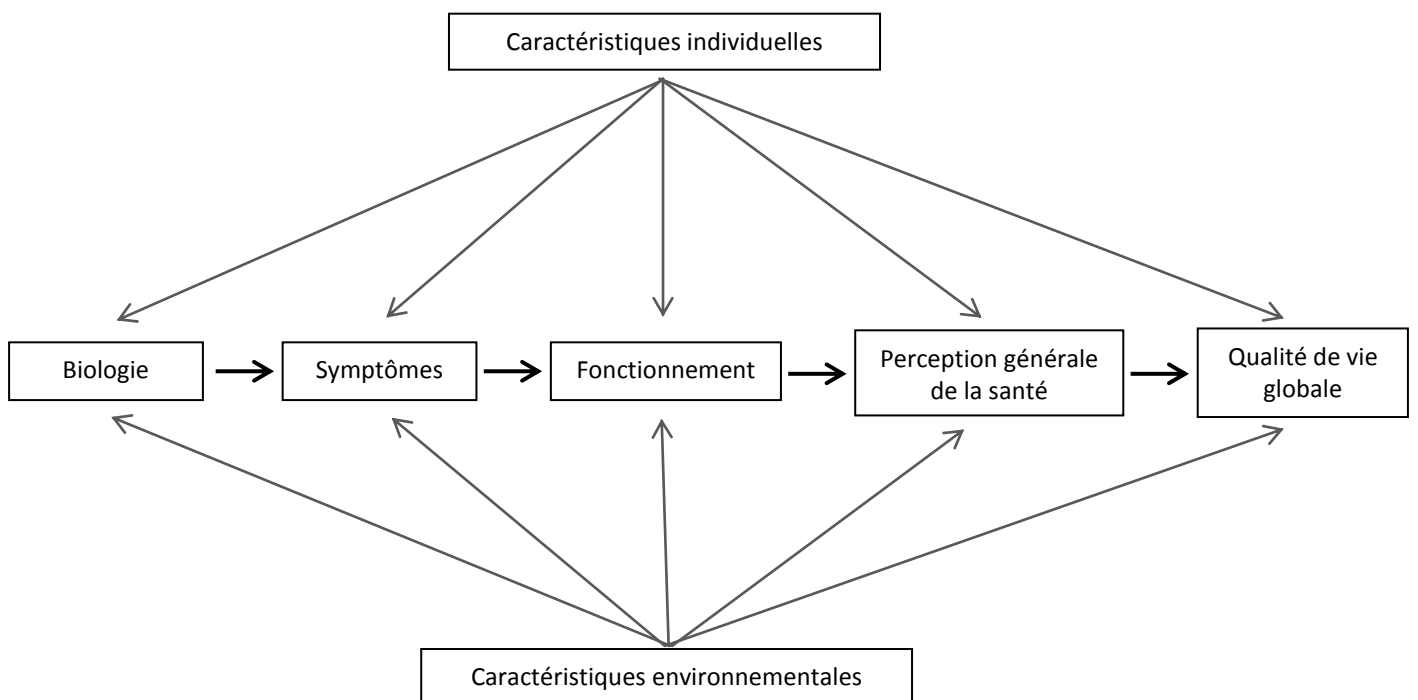
Contrairement aux modèles proposés antérieurement, que nous ne présenterons pas ici [16-19], celui-ci rassemble et met en relation deux aspects de la recherche : les sciences médicales et les sciences sociales.

Ce modèle constitué de 5 niveaux permet de visualiser l'enchaînement des facteurs biologiques, cliniques, physiques, psychologiques et sociaux qui modulent la qualité de vie (**figure 1**).



**Figure 1 :** Modèle conceptuel représentant les relations entre les différents éléments nécessaires à l'évaluation de la qualité de vie, d'après Wilson et Cleary [2].

D'autres équipes ont depuis travaillé sur ce modèle pour développer et préciser l'influence des facteurs environnementaux et individuels.



**Figure 2 :** Révision du modèle conceptuel de Wilson et Cleary, d'après Ferrans [20]

Le modèle révisé, présenté **figure 2**, ne spécifie plus quel est le facteur qui influe sur le niveau étudié. Ainsi, les facteurs psychologiques, les facteurs sociaux, les facteurs économiques, les motivations personnelles... sont supprimés pour laisser place à une généralisation. Le niveau biologie, qui n'était soumis à aucune influence extérieure, est à présent soumis aux influences des caractéristiques individuelles et environnementales. L'influence spécifique des facteurs non médicaux sur la qualité de vie globale est supprimée. Les facteurs non médicaux sont considérés comme faisant partie intégrante des facteurs individuels et environnementaux [20].

Cette *aggiornamento*, plus généraliste et plus simple, va permettre de proposer un cadre conceptuel pour développer des instruments de mesures méthodologiquement robustes car le modèle permet de mieux appréhender la notion de qualité de vie et ses éléments intrinsèques [2, 20].

## **5. Place de la qualité de vie en tant que critère de jugement en oncologie**

On appelle critère de jugement en oncologie, les critères permettant de conclure sur l'efficacité thérapeutique d'un traitement en oncologie.

Si on se réfère aux directives de la *Food and Drug Administration* (FDA), les critères de jugement principaux sont la survie globale (délai entre la date de diagnostic de la maladie et la date de décès) et les critères basés sur l'évaluation de la tumeur, à savoir le taux de réponse (proportion de patients pour lesquels le traitement a eu un effet significatif sur la tumeur), la survie sans progression (par exemple, délai entre une date de début de traitement et la date où la première évolution de la maladie est constatée ou le décès quelle qu'en soit la cause) et la survie sans maladie (délai pendant lequel le patient ne présente plus de maladie ou le décès quelle qu'en soit la cause) [21-25].

Si la FDA propose également des critères basés sur l'évaluation des symptômes (tolérance ou qualité de vie) comme critères de jugement potentiels, elle ne les présente pas comme des critères principaux (hormis, pour les essais de phase I, la tolérance est l'objectif principal puisqu'il s'agit de définir des doses limitantes toxiques).

Bien que critère secondaire, la tolérance, demeure un critère de jugement majeur pour les essais cliniques de phase II et III. Pour certains essais de phase II, le critère de jugement principal est d'ailleurs composé des deux paramètres que sont l'efficacité d'un traitement et la tolérance [26].

Couramment, si la FDA reconnaît l'importance de prendre en compte le point de vue du patient dans l'évaluation des soins et de la santé [27], elle met l'accent sur plusieurs aspects limitants de cette démarche :

- le recueil parfois malaisé des données ;
- la fréquence des données manquantes ;
- le manque de validité des instruments permettant de la mesurer ;
- la nécessité que la personne interrogée ou la personne qui recueille les données ne transforme pas la réalité ;
- la difficulté d'interprétation des résultats et de conclusion [22].

Néanmoins, de nombreux auteurs promeuvent la qualité de vie en tant que critère de jugement dans les essais cliniques et particulièrement en oncologie [10, 27, 28].

Ainsi, dès 1989 l'évaluation de la qualité de vie en tant que critère de jugement dans les essais de phase III randomisés a été proposée, après qu'elle eut permis de mettre en évidence une différence entre deux traitements alors que les critères habituels ne le permettaient pas [29].

A partir de cette première approche, de nombreux articles sont venus encourager l'utilisation de la qualité de vie comme critère de jugement. Les raisons sont nombreuses, nous proposons ici de citer les plus importantes [7, 9, 10, 15, 27, 30, 31] :

- il est nécessaire de prendre en compte le ressenti du patient dans la prise en charge de sa maladie – en effet, le patient est le meilleur témoin pour constater une modification de sa santé ;
- les critères standards ne tiennent pas compte des aspects physiques, psychologiques, sociaux et environnementaux dans l'évaluation d'un traitement – un nouveau traitement, une nouvelle fréquence d'administration... ne doivent pas détériorer la qualité de vie d'un patient ;
- une meilleure qualité de vie pourrait influencer la durée de survie des patients ;
- un meilleur contrôle de la chronicité de la maladie est actuellement en développement, mais se pose la question des conditions de vie pour le patient ;
- les nouvelles molécules proposées ne permettent pas forcément un gain de survie, mais peuvent permettre, pour une même efficacité, d'améliorer la qualité de vie du patient ;
- la qualité de vie est utilisée dans le cadre des analyses médico-économiques de type coût-utilité.

Par conséquent, il est désormais justifié d'utiliser la qualité de vie comme critère de jugement principal dans des études :

- cherchant à montrer l'équivalence ou la non-infériorité en terme d'efficacité d'un traitement ;
- comparant des traitements de chimiothérapie adjuvante ;

- cherchant à mettre en évidence une très faible augmentation en terme de survie du patient.

La qualité de vie comme critère de jugement principal s'impose aussi, d'évidence, dans un contexte de fin de vie où les possibilités thérapeutiques sont limitées, où les symptômes peuvent être extrêmement invalidants [32, 33]. La question de poursuivre un traitement parfois terriblement toxique, pour un gain potentiellement faible en terme de survie, se pose. D'après une étude récente, 7% des patients atteints d'un cancer ne sont pas prêts à accepter les toxicités considérables d'une chimiothérapie pour un gain de survie de 1 an [34].

## Chapitre 2 : Les instruments de mesure de la qualité de vie

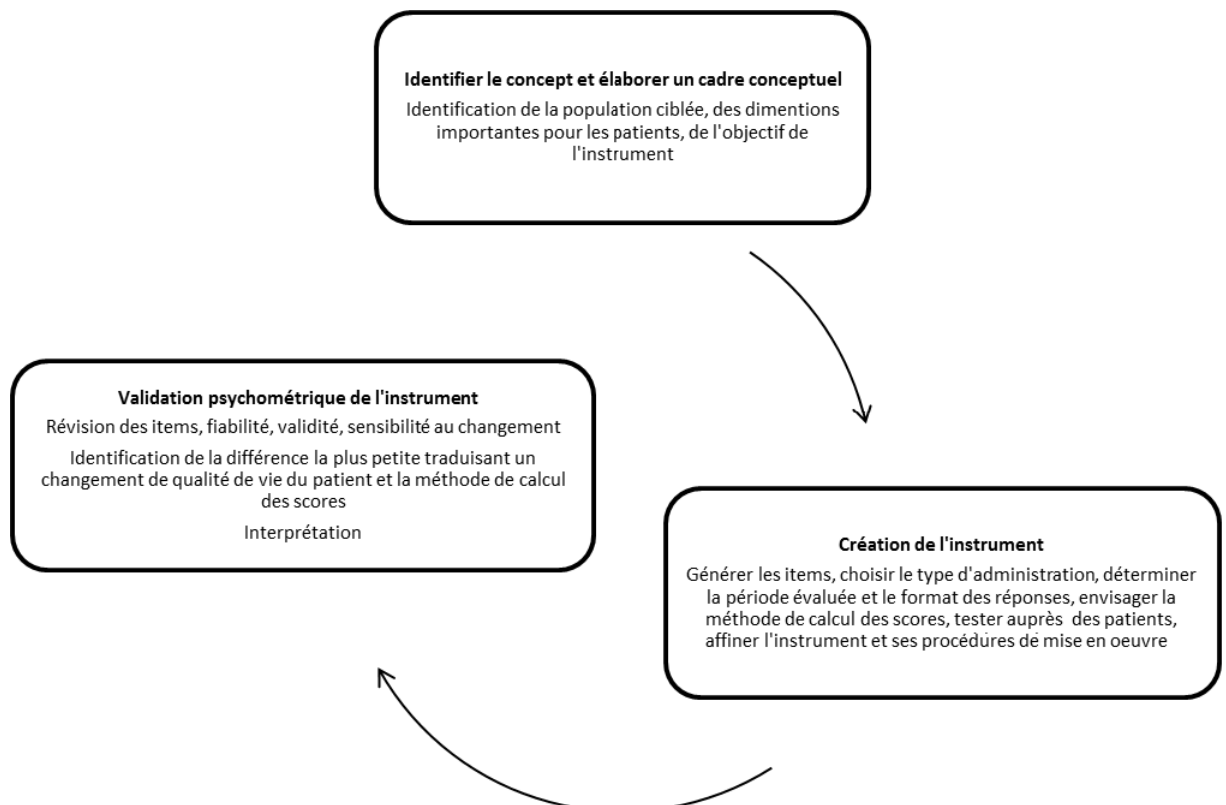
« Mesurer, c'est affecter un score à un objet ; [...]. Le processus de mesure requiert donc l'utilisation d'une échelle (ou d'un instrument) de mesure » [35]

Le développement d'un instrument permettant la mesure de la qualité de vie des patients suit un processus long et complexe qui nécessite de respecter certaines étapes pour construire un instrument performant.

Les instruments de mesure se présentent, le plus souvent, sous forme de questionnaires.

Dans un premier temps, il est important d'identifier clairement ses objectifs, à savoir la population à laquelle l'instrument de mesure sera destiné, et les concepts que l'on souhaite évaluer (cf. chapitre 1, § 3).

Une fois ces objectifs déterminés, le processus de création du questionnaire peut-être lancé. Cette partie de la revue de la littérature propose de mettre en exergue des éléments de réflexion permettant de développer un instrument de mesure de la qualité de vie et les étapes de validation essentielles.



**Figure 3 :** Processus de développement d'un instrument de mesure de la qualité de vie, inspiré du schéma proposé par la FDA [36]



## 1. Instrument : générique ou spécifique

Selon les objectifs de recherche, il est intéressant de disposer d'outils permettant soit d'évaluer la qualité de vie de façon globale et comparer les populations entre elles sans tenir compte de leurs spécificités, soit permettant d'évaluer la qualité de vie de façon plus ciblée et de prendre en considération toutes les subtilités qu'engendre tel ou tel type de maladie, tel état de santé, etc.

Ainsi trouve-t-on dans la littérature deux grandes familles de questionnaires : les questionnaires génériques et les questionnaires spécifiques.

Les questionnaires génériques sont utilisables et utilisés pour évaluer aussi bien différents types de populations que différents types de pathologies. Le *Medical Outcome Study -36 items Short-form* (SF-36) est un de ces questionnaires [37, 38]. Composé de 36 items, ce questionnaire permet d'évaluer la qualité de vie de tout type de population à travers huit domaines (activité physique, limites liées à l'activité physique, douleurs physiques, santé perçue, vitalité, vie et relations avec les autres, santé psychique, limites liées à l'état psychique).

Les questionnaires spécifiques, quant à eux, ciblent soit une population précise, soit une pathologie bien définie, soit un symptôme particulier.

Parmi les questionnaires spécifiques d'une population, citons les questionnaires généraux spécifiques d'une patientèle cancéreuse tel le *Quality of life Questionnaire-Core 30* (QLQ-C30) de l'*European Organization for Research and Treatment of Cancer* (EORTC) qui évalue six domaines (physique, psychologique, social, cognitif, fonctionnement personnel, problèmes financiers) et neuf symptômes (fatigue, nausées, vomissements, douleur, dyspnée, insomnie, appétit, constipation, diarrhée) [39], ou tel le *Functionnal Assessment of Cancer Therapy –General* (FACT-G) qui explore quatre dimensions (bien-être physique, bien-être social, bien-être fonctionnel, bien-être émotionnel) [6].

Parmi les questionnaires spécifiques d'une pathologie, citons à titre d'exemple certains développés par l'EORTC [40], dédiés à différentes localisations cancéreuses : le H&N35 pour les cancers de la tête et du cou [41], le BR23 pour les cancers du sein [42], le OV28 pour une localisation ovarienne [43].

Comme exemple de questionnaire spécifique d'un symptôme, citons le *Functionnal Assessment of Cancer Therapy – Anemia* (FACT-An), qui a été développé spécifiquement pour les patients anémiés. Il est composé du FACT-G et de 20 questions supplémentaires spécifiques à l'anémie [44].

## **2. Les différentes modalités d'évaluation**

### 2.1. Les entretiens structurés

Un entretien structuré est un entretien où la personne en charge de l'évaluation interroge le patient sur l'ensemble de sa vie. C'est l'évaluateur qui dirige la conversation, mais laisse le patient parler [45].

Si ce type d'évaluation permet d'un point de vue clinique d'évaluer la qualité de vie du patient dans son ensemble, les données récoltées sont, quant à elles, difficilement utilisables et transformables en mesure, parce qu'il s'agit de « réponses –discours », donc non standardisées, et non de choix dans une liste prédéfinie.

C'est une méthode dite qualitative, qui cherche à comprendre l'autre.

### 2.2. Les auto-questionnaires

Un auto-questionnaire est un questionnaire édité sur papier ou disponible sur tablette tactile et laissé au patient pour qu'il complète seul chacun des items.

D'après la définition de la qualité de vie, quel qu'en soit l'auteur, l'utilisation des auto-questionnaires est à privilégier. En effet, qui mieux que le patient peut indiquer dans quelle mesure des difficultés, qu'elles soient physiques, psychologiques ou sociales affectent son bien-être, et avec quelle intensité ?

L'auto-questionnaire est un questionnaire permettant d'évaluer différentes dimensions participant au bien-être du patient.

Il est cependant courant que le patient soit aidé par un tiers (un proxy) pour lire et compléter le questionnaire, surtout si son état de santé ne lui permet pas de le faire seul. Il est admis que si le patient est aidé pour répondre, les réponses peuvent ne pas être celles qu'il aurait indiquées s'il avait été seul [46]. Dans cette optique, il est important que la personne qui aide le patient lors de la passation des questionnaires ne soit ni le médecin référent du patient, ni un proche du patient [47]. Afin de tenir compte de cette potentielle source de variabilité, il est nécessaire, lors de la récupération des données, d'indiquer quel est le rôle de la personne qui aide le patient. Cette information pourra alors être utilisée au cours de l'analyse des données.

### 2.3. Les questionnaires sous forme d'entretien

Ce sont des questionnaires qui imposent, tout particulièrement, à l'évaluateur de suivre scrupuleusement les consignes des auteurs.

L'évaluateur ne doit ni influencer, ni conseiller, ni commenter les réponses données par le patient. Il se doit donc d'observer la plus grande neutralité.

Ce type de questionnaire est principalement utilisé en psychiatrie [48] ou en pédiatrie [49-51]. Il n'est pas rare qu'un questionnaire puisse être à la fois présenté sous forme d'entretien et d'auto-questionnaire. Ainsi par exemple, le guide d'utilisation du FACT-G précise les conditions d'utilisation en cas de passation par un membre de l'équipe médicale [44]. Il en est de même pour le *Brief Fatigue Inventory* (BFI) [52] ou le *Missoula Vitas Quality Of Life Index* (MVQOLI) [53].

## 2.4. Les hétéro-questionnaires

Un hétéro-questionnaire est un questionnaire où la qualité de vie du patient est évaluée par une autre personne que le patient lui-même. L'évaluateur peut être un membre de l'équipe médicale ou un proche du patient.

En oncologie, classiquement, les médecins utilisent, probablement sans le savoir, un hétéro-questionnaire pour évaluer la santé globale de leur patient, il s'agit du *Eastern Cooperative Oncology Group performance status scale* (ECOG Performance Status) [54].

En soins palliatifs, il existe le *Performance Palliative Scale* (PPS), le médecin indique la mobilité du patient, l'activité du patient et l'intensité de la maladie, l'autonomie du patient pour ses soins personnels, son type d'alimentation et son niveau de conscience. Selon ces réponses, le médecin va obtenir un score en pourcents qui va lui indiquer le niveau de performance de son patient. Ce score lui permettra d'avoir une vision globale de l'état de santé du patient mais pourra également lui servir de facteur pronostique quant au décès [55-57].

Les hétéro-questionnaires, sont également employés lorsqu'il s'agit d'évaluer la qualité des soins à travers la qualité de vie des patients en fin de vie. Après le décès de leur proche, les membres de la famille sont parfois interrogés sur leurs ressentis concernant la fin de vie du membre de leur famille décédé [58].

## **3. Le format des réponses**

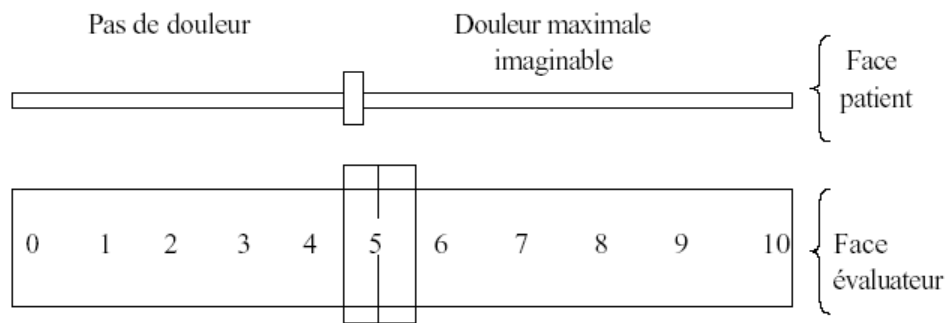
Il existe plusieurs formats de réponses :

### 3.1. Les échelles visuelles analogiques (EVA)

Il s'agit d'une échelle graduée de 0 à 10 permettant d'évaluer l'importance prise par l'évènement. Le zéro étant l'absence de l'évènement, le 10 l'importance la plus intense imaginable [36].

Le patient indique une marque sur l'échelle pour exprimer le niveau d'importance de l'évènement.

A l'origine, cette échelle était utilisée uniquement pour évaluer la douleur. Il existe des réglettes où il suffit de faire avancer un curseur pour appliquer la marque sur l'échelle.



**Figure 3 :** Echelle visuelle Analogique, d'après <http://www.soins-infirmiers.com/douleur.php>

### 3.2. Les échelles de type Likert

Il s'agit d'une échelle composée de propositions de réponse avec un degré d'accord à la proposition qui est croissant (réponse ordinale). Le plus souvent, l'échelle est utilisée sous une forme composée de cinq ou sept réponses possibles. Lorsque le nombre de réponses proposées est impair, la valeur centrale permet à la personne interrogée d'exprimer sa neutralité par rapport à la proposition.

- |  |  |  |  |   |  |   |
|--|--|--|--|---|--|---|
| <input type="radio"/> d'accord<br>agree                      | <input type="radio"/> pas d'accord<br>disagree           |  |  |   |  |   |
| <input type="radio"/> bonne<br>good                          | <input type="radio"/> moyenne<br>fair                    | <input type="radio"/> pauvre<br>poor           |  |   |  |   |
| <input type="radio"/> souvent<br>often                       | <input type="radio"/> parfois<br>sometimes               | <input type="radio"/> rarement<br>seldom       | <input type="radio"/> jamais<br>never      |   |  |   |
| <input type="radio"/> tout à fait d'accord<br>strongly agree | <input type="radio"/> d'accord<br>agree                  | <input type="radio"/> undecided<br>indécis     | <input type="radio"/> d'accord<br>disagree | <input type="radio"/> pas du tout d'accord<br>strongly disagree |  |   |
| <input type="radio"/> toujours<br>always                     | <input type="radio"/> souvent<br>very frequently         | <input type="radio"/> parfois<br>occasionally  | <input type="radio"/> rarement<br>rarely   | <input type="radio"/> jamais<br>never                           |  |   |
| <input type="radio"/> sûrement<br>definitely                 | <input type="radio"/> très probablement<br>very probably | <input type="radio"/> probablement<br>probably | <input type="radio"/> possible<br>possibly | <input type="radio"/> probablement pas<br>probably not          | <input type="radio"/> sûrement pas<br>definitively not |   |
| <input type="radio"/> extrêmement utile<br>extremely useful  | <input type="radio"/> très utile<br>quite useful         | <input type="radio"/> utile<br>slightly useful | <input type="radio"/> neutre<br>neutral    | <input type="radio"/> inutile<br>slightly useless               | <input type="radio"/> très inutile<br>quite useless    | <input type="radio"/> parfaitement inutile<br>extremely useless |

**Figure 4 :** Les échelles de type Likert, d'après : <http://www.definition-qualite.com/echelle-de-likert.htm>

La passation d'un questionnaire qui repose sur des échelles de Likert à cinq ou sept points apparaît comme étant plus simple d'utilisation, à la fois par le patient et par le personnel médical, car il facilite l'interprétation des données tout en offrant une variabilité de réponses [36, 59].

### 3.3. Les échelles à extrémités

Il s'agit d'une échelle où les extrémités sont spécifiées, mais pas les possibilités intermédiaires. Le patient coche une case pour indiquer son niveau. C'est donc lui qui interprète la valeur des possibilités intermédiaires. Le principe est le même que pour l'EVA [36].



**Figure 5 :** Les échelles à extrémités

### 3.4. Les listes récapitulatives

Il s'agit de liste d'évènements avec des propositions de réponses réduites, Oui, Non, Ne sait pas [36].

## **4. La mesure obtenue**

Les réponses du patient au questionnaire sont des données qualitatives, il est donc nécessaire d'effectuer une transformation mathématique afin d'obtenir un ou des scores permettant de mesurer la qualité de vie du domaine évalué ou la qualité de vie globale. C'est lors de la construction du questionnaire que cette mesure est définie.

Le principe de base est d'attribuer une valeur, un score à chaque item suivant l'importance que l'item représente pour le patient. En fonction de son caractère positif ou négatif le score peut être attribué de façon croissante ou décroissante (*reverse coding*).

Très souvent, les scores des différents domaines évalués ou score globaux, sont obtenus en sommant les valeurs des différents items, c'est le cas pour le questionnaire QLQ-C30 de L'EORTC, qui est un questionnaire très fréquemment employé en oncologie. Après normalisation, les scores varient de 0 (pire) à 100 (meilleure).

D'autres questionnaires reposent sur une pondération, c'est le cas pour le questionnaire *Missoula-VITAS Quality of Life Index* par exemple [53].

## 5. Les qualités psychométriques des instruments de mesure de la qualité de vie

« La qualité d'une mesure est déterminée par la réponse à deux questions : Que vaut la mesure ? Quel sens donner à la mesure ? Dans le champ des mesures objectives la réponse à la seconde question est grandement facilitée par l'existence de modèles physiopathologiques solides. Ces modèles n'existent pas le plus souvent dans le champ des mesures subjectives, il faut donc être particulièrement prudent, en pratique, avant d'accepter que tel instrument mesure bien ce qui est annoncé dans son intitulé... Ce point est particulièrement vrai en ce qui concerne les mesures dites de qualité de vie »[60].

Avant de pouvoir être utilisés dans les unités de soins, les questionnaires doivent faire l'objet de validations psychométriques. Cette validation est composée de plusieurs étapes nécessaires à la pertinence des scores obtenus et à l'interprétation qui en sera faite.

### 5.1. La validité du fond

L'étape de validation du fond (*face validity* et *content validity*) est un des tous premiers critères à vérifier lors de la construction d'un nouvel outil de mesure. Cette étape doit permettre de s'assurer (1) que les questions sont compréhensibles, sans double sens ou ambiguïtés, (2) que les questions correspondent bien aux domaines qu'elles abordent et (3) que les domaines évalués correspondent bien à la population à laquelle le questionnaire est destiné.

Ce sont, en général, des experts qui vérifient ces trois points, médecins, méthodologistes, infirmiers, psychologues, etc..

Pour le premier point (*face validity*), il est important de faire relire les questions établies par des patients, afin d'avoir leur vision, leur compréhension du texte [36].

Pour les points 2 et 3 (*content validity*), une revue de la littérature ou une interrogation directe des patients quant aux domaines qui les concernent de façon plus prégnante doivent permettre, après une sélection, d'identifier les domaines incontournables que devra évaluer l'instrument [35, 61].

La sélection, peut se faire dans un premier temps suivant la fréquence de représentation de chaque domaine cité, puis dans un second temps suivant une analyse factorielle des correspondances des données [62-64].

L'analyse factorielle des correspondances permet une représentation graphique de la répartition des variables étudiées et donc de définir les distances qui les séparent les unes des autres et ainsi constituer des regroupements d'items.



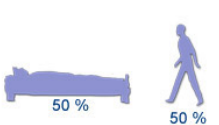
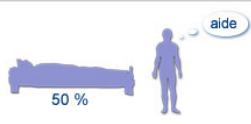


## 5.2. La validité du critère

La validation du critère (*criterion validity*) doit permettre de s'assurer que le nouvel outil dans son ensemble mesure ce qu'il est censé mesurer. Concrètement, cela consiste en l'évaluation des relations qui existent entre le nouvel outil de mesure et un instrument de référence qui possède la même structure, c'est-à-dire qui évalue les mêmes domaines de la même population cible. Statistiquement, cela consiste en la réalisation de tests de corrélation (convergence ou divergence).

Or, si l'instrument de référence n'existe pas (ce qui est souvent le cas), il est nécessaire d'utiliser plusieurs critères pour valider la construction du nouvel outil de mesure (*construct validity*). Il s'agit alors, d'évaluer la corrélation qui existe entre le nouvel outil et des critères qui explorent les mêmes dimensions. Le but étant de conclure que l'outil nouvellement créé va dans le même sens que les critères [36, 39, 61].

Par exemple, le résultat obtenu pour le domaine physique doit être corrélé, aller dans le même sens, que l'*ECOG Performance Status Scale*.

L'*ECOG Performance Status Scale* est gradé comme suit :

Capable d'une activité identique à celle précédant la maladie sans aucune restriction.		0
Activité physique diminuée mais ambulatoire et capable de mener un travail.		1
Ambulatoire et capable de prendre soin de soi-même, incapable de travailler. Alité moins de 50 % de son temps.		2
Capable seulement de quelques soins. Alité ou en chaise plus de 50 % du temps.		3
Incapable de prendre soins de soi-même. Alité ou en chaise en permanence.		4
Décès		5

**Figure 6** : Cotation de l'*ECOG Performance Status Scale*, d'après : <http://www.cancer-et-denutrition.com/cancer-et-denutrition/42.htm>

Ainsi on va chercher à calculer le coefficient de corrélation qui relie le score du nouvel outil et l'*ECOG Performance Status Scale*.

En utilisant le coefficient de corrélation de Spearman, le coefficient se calcule de la manière suivante (formule approchée, ne tenant pas compte du nombre d'ex-aequos au sein du groupe) [65] :

$$r_{Spearman} = 1 - 6 \sum (x_i - y_i)^2 / n(n^2 - 1)$$

Les valeurs  $x_i$  et  $y_i$  correspondent aux rangs des observations,  $x$  étant l'ECOG (*Performance Status Scale*) et  $y$  le score obtenu avec le nouvel instrument de mesure pour l'individu  $i$ ,  $n$  est l'effectif de la population étudiée.

Si le coefficient obtenu est proche de 1, alors le nouvel outil est convergent avec le critère référent (*ECOG Performance Status Scale*). Si le coefficient obtenu est proche de -1, le nouvel outil est divergent et s'il est proche de 0, l'outil n'est pas corrélé au critère de référence, il ne mesure donc pas ce pour quoi il était créé.

Le test de nullité du coefficient de corrélation de Spearman suit une loi de Student à  $(n-2)$  degré de liberté.

### 5.3. La fiabilité

Sous l'aspect de fiabilité (*reliability*), se cachent deux notions : la cohérence et la reproductibilité.

#### **La cohérence interne**

Evaluer la cohérence (cohérence interne), c'est s'assurer que les questions qui constituent l'évaluation d'un domaine soient homogènes entre elles.

Pour évaluer la cohérence interne, on utilise, classiquement, l'alpha de Cronbach.

Cette statistique permet de quantifier la force des corrélations entre questions. En pratique, celui-ci doit être supérieur à 0,7 (sur un maximum de 1) pour permettre de conclure à une bonne homogénéité des questions dans un domaine si l'on veut utiliser l'outil pour comparer des populations.

Pour une application quotidienne de l'outil, donc utilisé en pratique pour l'évaluation d'un seul et unique patient à la fois, il est nécessaire d'obtenir un alpha de Cronbach d'au moins 0.90 [66].

L'alpha de Cronbach, se calcule comme suit [67] :

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum \sigma_i^2}{\sigma_T^2} \right)$$

Où  $k$  est le nombre de questions composant le domaine étudié,



$\sigma_i^2$  est la variance de la  $i^{\text{ème}}$  question et  $\sigma_T^2$  est la variance totale formée par l'ensemble des questions

Attention, si le score n'est pas un score additionnel, mais pondéré, il est nécessaire de pondérer également la question avant le calcul de la variance.

Récemment, de nouvelles méthodes telles que la théorie du trait latent ou théorie des réponses à l'item (modèles IRT, modèle de Rasch) [68-71] ont été développés pour palier à l'approche des scores comme des variables quantitatives continues. En effet, le calcul de l'alpha de Cronbach n'est pas critiquable lorsque l'échelle utilisée est une EVA classique (c'est-à-dire où l'on vient mesurer la distance entre le zéro et la marque faite par le patient), il s'agit bel et bien d'une variable quantitative continue. En revanche il est discutable d'utiliser le coefficient de Cronbach, lorsque les réponses du patient sont issues des propositions d'une échelle de Likert. Dans ce contexte, le score attribué est une variable quantitative discrète.

Les modèles IRT s'appuient sur une relation entre la probabilité de répondre à une des modalités d'un item (parmi les autres) et une variable sous-jacente non évaluable (le trait latent). Le trait latent explique la réponse donné à chaque item.

« Le trait latent est une particularité humaine [...] présente chez chacun et qui se manifeste à des degrés divers selon les individus » [68].

Le modèle de Rasch fait partie des modèles IRT. Il est le plus simple à utiliser. C'est le modèle que nous avons choisi de présenter ici.

Chaque item est caractérisé par un paramètre unique appelé le paramètre de difficulté, il est utilisé lorsque les réponses aux items prennent des modalités dichotomiques (positif / négatif, les modalités les plus favorables sont définies comme positives, les plus défavorables comme négatives).

La probabilité qu'un patient  $i$  réponde  $x_{ij}$  à l'item  $j$  est modélisée par le modèle logistique à deux paramètres :

$$\Pr(X_{ij} = x_i \mid \theta_i, \delta_j) = \frac{\exp(x_{ij}(\theta_i - \delta_j))}{1 + \exp(\theta_i - \delta_j)}$$

Le trait latent est noté  $\theta_i$  pour le patient  $i$  et la « difficulté » de réponse à l'item  $j$ , noté  $\delta_j$ , la variable aléatoire  $\theta_i$  suit une distribution normale.

La probabilité que le patient réponde positivement à l'item  $j$  est donc :

$$\Pr(X_{ij} = 1 \mid \theta_i, \delta_j) = \frac{\exp(\theta_i - \delta_j)}{1 + \exp(\theta_i - \delta_j)}$$

L'allure de la courbe représentant la distribution de probabilité des réponses positives de chaque item est en forme de S. Le seul paramètre à estimer dans ce modèle est l'abscisse du point d'inflexion de la courbe.

Le modèle de Rash, est un modèle assez simple à interpréter, la valeur du paramètre  $\delta_j$  représente la valeur du trait latent qu'il faut atteindre pour avoir 50% de chance de répondre positivement à l'item.

### **La reproductibilité**

Evaluer la reproductibilité, c'est s'assurer que le questionnaire, s'il est complété par un même patient 2 fois, dans un intervalle de temps assurant la stabilité de sa qualité de vie, permet de reproduire des résultats similaires. En d'autres termes, il s'agit de voir si l'évaluation que va permettre le questionnaire est homogène s'il n'y a pas de modification de l'état de santé du patient même si le patient n'a pas indiqué exactement les mêmes réponses que lors du premier remplissage.

C'est ce que l'on cherche à vérifier avec la méthode dite de « test-retest » [72]. Cette méthode consiste à faire compléter, dans les mêmes conditions, le questionnaire évalué par un patient deux fois dans un intervalle de temps fixé (3 à 15 jours). Ensuite, il est nécessaire de comparer statistiquement les résultats obtenus entre les deux passations des questionnaires.

Dans cette optique, deux méthodes peuvent être employées :

Pour les scores, le calcul du Coefficient de Corrélation Intra-classe (CCI) [73, 74]. Ce coefficient est issu du travail de Fisher réalisé dans le cadre du développement des modèles aléatoires de l'analyse de la variance. Dans ce contexte, la formule obtenue permettant de calculer le CCI est :

$$CCI = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\epsilon}^2}$$

Où  $\sigma_{\alpha}^2$  est la variance de l'effet aléatoire du patient et  $\sigma_{\epsilon}^2$  est la variance de l'erreur.

Plus le CCI est proche de 1 plus la reproductibilité est bonne. Un CCI > 0.70 est généralement considéré comme le seuil permettant d'indiquer que la reproductibilité est bonne [75].

Pour les items, le calcul du Coefficient Kappa de Cohen [74, 76]:

$$kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}$$

Où Pr(a) est la proportion de correspondances observées et Pr(e) est la proportion d'accords aléatoires.

C'est-à-dire

$$\Pr(a) = \frac{1}{n} \sum n_{ii}$$

$$\Pr(e) = \frac{1}{n^2} \sum n_{i,n_i}$$

Avec les effectifs n présenté ci-dessous :

	Passation 2					Total
	Questionnaire	Réponse 1	Réponse 2	... Réponse r		
Passation 1	Réponse 1	$n_{11}$	$n_{12}$	...	$n_{1r}$	$n_{1.}$
	Réponse 2	$n_{21}$	$n_{22}$	...	$n_{2r}$	$n_{2.}$
	...					
	Réponse r	$n_{r1}$	$n_{r2}$	...	$n_{rr}$	$n_{r.}$
	Total	$n_{.1}$	$n_{.2}$	...	$n_{.r}$	$n$

Extrait de : [http://kappa.chez-alice.fr/Kappa\\_2juges\\_Def.htm](http://kappa.chez-alice.fr/Kappa_2juges_Def.htm)

Suivant la statistique obtenue, on interprète le résultat selon la classification proposée par Cohen [77] :

Correspondance	Kappa
Excellente	> 0.80
Bonne	]0.60 ; 0.80]
Modérée	]0.40 ; 0.60]
Médiocre	]0.20 ; 0.40]
Mauvaise	]0 ; 0.20 ]
Très mauvaise	≤ 0

#### 5.4. La sensibilité au changement

La notion de sensibilité au changement est apparue dans la littérature en 1985 [78]. Depuis, elle fait débat.

Certains auteurs la définissent comme un nouveau critère, d'autres, l'incluent dans la reproductibilité, d'autres encore la considèrent comme un critère de validité [79].

Toutefois, une entente existe pour dire qu'elle est nécessaire, voire essentielle à la validation d'un instrument de mesure de la qualité de vie.

Il n'y a pas de consensus sur la définition de sensibilité au changement. La conception du critère diffère selon les auteurs. « 25 définitions et 31 mesures ont été trouvées » [80].

D'après la revue de la littérature réalisée en 2010, nous pouvons distinguer 2 types de sensibilité au changement ou *responsivness* [81].

### ***La sensibilité au changement : Internal responsivness***

La difficulté, sur laquelle repose l'évaluation de la qualité de vie est l'interprétation que l'on peut faire de la mesure. Il est donc important de définir à partir de quelle mesure ou de quelle variation l'outil traduit une modification de l'état de santé ou la qualité de vie du patient.

Un outil est dit être sensible au changement (*internal responsivness*) s'il peut mettre en évidence un changement sans qu'il soit nécessairement cliniquement significatif [81].

Il existe plusieurs méthodes permettant de mettre en évidence la sensibilité au changement de l'outil, nous présenterons ici les plus souvent utilisées [78, 81] :

#### **La taille de l'effet (effect size) :**

$$T_{aill\grave{e}}E_{ffet} = \frac{M_1 - M_0}{ET(M_0)}$$

Où  $M_0$  la moyenne des scores de la passation n°1 et  $M_1$  la moyenne des scores de la passation n°2, ET est l'écart-type (l'écart à la moyenne)

Entre 0.2 et 0.5 l'effet est petit, entre 0.5 et 0.8 l'effet est moyen et au-delà de 0.8 l'effet est grand.

#### **La réponse moyenne normalisée (Standardized response mean) :**

$$R_{\acute{e}ponse}M_{oyenne}S_{tandardis\acute{e}e} = \frac{M_1 - M_0}{ET(M_1 - M_0)}$$

Où  $M_0$  la moyenne des scores de la passation n°1 et  $M_1$  la moyenne des scores de la passation n°2,  $ET(M_1 - M_0)$  est l'écart-type du changement observé (l'écart à la moyenne du changement)

Une mesure qui possède un haut niveau de variabilité dans son score aura une petite Réponse Moyenne Standardisée et donc une mauvaise sensibilité au changement, puisque entre 0.2 et 0.5 de l'effet est petit, entre 0.5 et 0.8 l'effet est moyen et au-delà de 0.8 l'effet est grand.

#### **Un test pour comparaison de moyenne pour série appariée (Paired t-test) :**

Les hypothèses testées sont les suivantes :

$$\begin{cases} H_0 : (M_1 - M_0) = 0 \\ H_1 : (M_1 - M_0) \neq 0 \end{cases}$$

La statistique de test t est calculée comme suit :

$$t = \frac{M_1 - M_0}{ET(M_1 - M_0)/\sqrt{n}}$$

Où  $M_0$  la moyenne des scores de la passation n°1 et  $M_1$  la moyenne des scores de la passation n°2,  $ET(M_1-M_0)$  est l'écart-type du changement observé (l'écart à la moyenne du changement) et  $n > 30$ .

La statistique de test, en valeur absolue, est ensuite comparée à la valeur seuil classiquement employée de 1.96 (alpha 5%). Si la statistique de test est supérieure alors on rejette l'hypothèse nulle, on peut donc conclure à une sensibilité au changement de l'outil testé.

### ***La pertinence : External responsiveness***

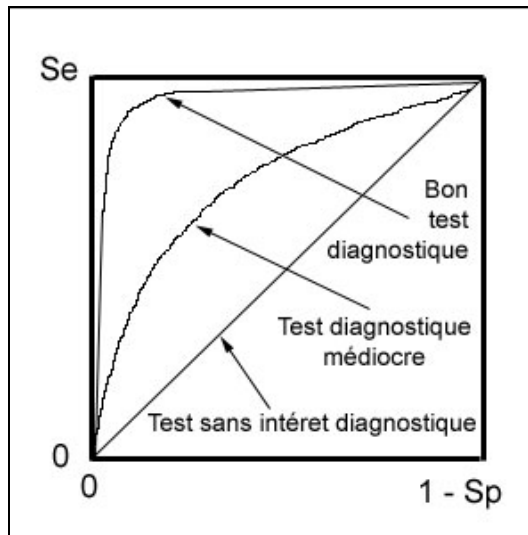
Un outil est pertinent (*external responsiveness*), s'il est capable de mettre en évidence un changement cliniquement significatif [15, 78, 81]. Il s'agit de mettre en relation l'outil évalué avec une donnée clinique permettant d'évaluer l'état de santé du patient.

Par exemple, le médecin estime que l'ECOG du patient est passé de 2 à 3 (cf. chapitre 2, §5.2), ce qui traduit une altération de l'état général du patient. Les questionnaires de qualité de vie qui auront été complétés lorsque le patient était ECOG 2 puis ECOG 3 devront présenter des scores différents.

Afin de mettre en relation le score et la référence clinique utilisée, plusieurs méthodes sont proposées :

#### **Evaluation des courbes ROC (Receiver Operating Characteristic)**

Cette méthode repose sur le calcul de la sensibilité, c'est à dire la probabilité de déterminer correctement si un patient présente bien un changement clinique et de la spécificité, la probabilité de déterminer correctement les patients pour lesquels il n'y a pas de changement clinique. Lorsque l'on trace sur un graphique la sensibilité et 1-spécificité pour chaque valeur observée de la mesure, on obtient la courbe ROC. Attention toutefois, pour calculer les probabilités il est nécessaire en premier lieu de dichotomiser la référence clinique (1 : changement, 0 : pas de changement). Plus l'aire sous la courbe est grande plus la capacité du score à déterminer un changement clinique est importante [81].



**Figure 7** : Courbe ROC, d'après [http://www.adscience.fr/uploads/ckfiles/files/html\\_files/StatEL/statel\\_courbe\\_ROC.htm](http://www.adscience.fr/uploads/ckfiles/files/html_files/StatEL/statel_courbe_ROC.htm)

### **Corrélation**

Le coefficient de corrélation de Pearson est souvent utilisé. Il se calcule comme suit [81] :

$$r_{xy} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right) \left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}$$

Où  $x_i$  et  $y_i$  correspondent respectivement aux scores obtenus aux premier et second temps de mesure de la qualité de vie.

Sous l'hypothèse de nullité du coefficient, le coefficient suit une loi de Student à  $n-2$  degré de liberté.

Attention toutefois à l'utilisation du coefficient de corrélation de Pearson, car il est sensible aux données aberrantes.

Utiliser le coefficient de Spearman permet de s'affranchir de la nécessité que la distribution du couple  $(x,y)$  suit une loi binormale pour de petits échantillons et de palier l'effet des données aberrantes (cf. chapitre 2 § 5.2).

### **Modèle de régression**

Réaliser une régression linéaire est une méthode alternative, ou une régression d'un autre type si la relation n'est pas linéaire [81].

Si on utilise un modèle linéaire :

$$D_{y_i} = a + bx_i + e_i$$

Où  $x_i$  est la différence entre les deux mesures du score pour l'individu  $i$ ,  
et  $y_i$  les valeurs de l'élément clinique (ECOG par exemple) pour le même individu,

$a$  est le changement moyen de  $y$  quand aucun changement n'est observé en  $X$ ,

$b$  représente la variation moyenne de  $y$  quand  $x$  change d'une unité.

Les  $e_i$  sont indépendants et  $e_i \sim N(0, \sigma^2)$ .

Quand  $b$  est proche de zéro, on peut conclure que le nouvel instrument n'est pas sensible au changement que  $X$  indique. Si  $b$  est grand,  $X$  et  $Y$  (la nouvelle mesure et le référent clinique) sont associés.

### ***La Différence Minimale Cliniquement Importante(DMCI)***

Afin d'utiliser la qualité de vie comme critère de jugement, il est nécessaire de mettre en évidence la différence minimale cliniquement importante (DMCI) [27, 30, 82, 83].

En effet, l'une des principales critiques développées à l'encontre de l'évaluation de la qualité de vie, et de son utilisation en tant que critère de jugement, est la difficulté d'interprétation qu'elle engendre.

La DMCI est définie par « la plus petite différence d'un score, d'une dimension évaluée, perçue par le patient comme bénéfique qui justifierait en l'absence d'effets indésirables et de coût excessif un changement dans la prise en charge du patient » [59]. D'un point de vue sociétal, la DMCI peut être perçue par le degré de changement qui stimulerait un clinicien, ou une proportion significative de cliniciens pour envisager une intervention dans une situation donnée [84].

Elle est issue de l'évaluation de la sensibilité au changement.

Dans un contexte de recherche clinique, il est classique de chercher à classer les patients selon leur état de santé, la recherche de la DMCI doit permettre identiquement de définir la composition de groupes sur la base de scores obtenus à une évaluation de la qualité de vie. Par exemple, Osoba a démontré que pour le QLQ-C30 de l'EORTC, une différence moyenne de 5 à 10 points entre les scores indique une modification mineure, de 10 à 20 une modification modérée et une différence de plus de 20 points une importante modification [84].

Afin de déterminer la DMCI, il est conseillé d'utiliser plusieurs méthodes et de confronter les résultats. C'est la méthode de la triangulation [30].

- L'utilisation d'une question générale « d'ancrage » qui interroge le patient sur son appréciation d'une modification sur le domaine évalué (*anchor*). La réponse est proposée sur la forme d'une échelle de Likert. Van Kampen propose une échelle de Likert en 7 points : complètement récupéré, beaucoup amélioré, légèrement amélioré, inchangé, légèrement moins bon, beaucoup moins bon, pire que jamais.

Les scores moyens sont calculés pour chacun des groupes constitués suivant le nombre de possibilité de l'échelle proposée (ici 7). La DMCI est définie comme la différence moyenne entre les deux scores des patients qui ont indiqué avoir eu une légère modification (bonne ou mauvaise) [83].

- Le calcul de l'Index de changement fiable [78, 85] :

$$I_{ndex\ Change\ Fiable} = \frac{(M_1 - M_0)}{1.96 \times ET_{change}}$$

$$ET_{change} = \sqrt{2\sigma_{pp}^2 + \sigma_{p \times T}^2}$$

Où  $\sigma_{pp}^2$  est la variabilité due à l'erreur de mesure des deux remplissages (1 et 2).

$\sigma_{p \times T}^2$  est la variabilité due à la différence entre les patients

- Le calcul de l'erreur standard de mesure définie par [86]:

$$E_{rreur\ Standard\ Mesure} = \sqrt{\sigma_x(1 - r_{xx})}$$

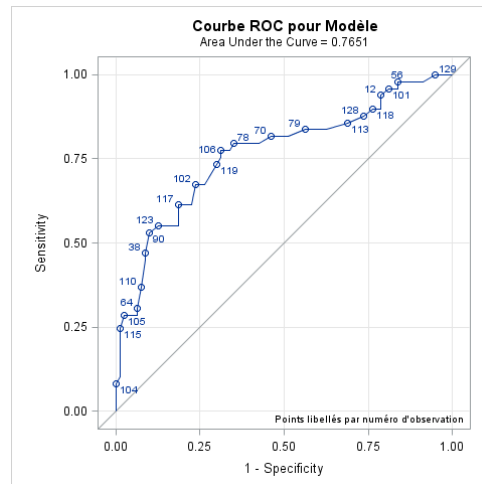
Où  $\sigma_x$  représente la variabilité du score sur la population,

Et  $r_{xx}$  le coefficient de fiabilité du score.

- L'utilisation des courbes ROC [87] :

Il s'agit d'identifier sur la courbe ROC le(s) point(s) sur la courbe qui se situent le plus près du coin en haut à gauche du cadre. La méthodologie pour tracer la courbe ROC est présentée au point ci-dessus 5.4.





**Figure 8 :** Identification des valeurs seuil, exemple de courbe ROC

La recherche de la DMCI, doit être bien documentée, et spécifique pour chaque population étudiée.

## 6. L'adaptation d'un questionnaire dans une autre langue

La spécificité de chaque culture fait qu'on ne peut pas se contenter de traduire un instrument développé dans une langue pour évaluer la qualité de vie de patients utilisant une autre langue.

La méthodologie est stricte, un instrument permettant de mesurer la qualité de vie doit être culturellement adapté et validé.

### 6.1. L'adaptation culturelle

L'adaptation culturelle peut être effectuée selon 2 méthodes :

#### **La traduction –rétro traduction (backward-forward translation)**

Cette méthode, peut être présentée en quatre étapes :

- (1) De façon indépendante, deux traductions des items du questionnaire source, sont réalisées, dans la langue du pays où le questionnaire doit être adapté, par deux personnes natives de ce pays mais parlant couramment la langue du pays où le questionnaire a été développé.

Un coordinateur, bilingue, va s'assurer qu'il n'y a pas de différences entre les deux traductions. Si des divergences apparaissent, le coordinateur, après discussions avec les traducteurs, homogénéisera la traduction.

- (2) De façon indépendante, deux traductions, de la traduction obtenue (étape 1), sont réalisées, dans la langue du pays d'origine du questionnaire, par deux personnes natives du pays d'origine du questionnaire mais parlant couramment la langue dans laquelle la traduction a été effectuée. Comme lors de la première étape, le coordinateur effectuera l'homogénéisation. C'est la rétro traduction.
- (3) Le questionnaire obtenu est proposé à des patients (10 à 30) [88-90], représentatif de la population vers laquelle est destiné le questionnaire. Chacun sera ensuite vu en entretien pour connaître : la difficulté de réponse, la compréhension, s'il y a confusion, les mots ou les phrases perturbantes.
- (4) Suivant ce test, le questionnaire est ensuite réévalué et modifié.

### **La confrontation de traductions indépendantes (Dual Panel Approach)**

Elle se décompose en trois étapes :

- (1) La traduction : quatre experts bilingues (linguistes, psychologues, méthodologistes...), au minimum, effectuent, de façon indépendante, une traduction du questionnaire d'origine dans la langue dans laquelle il doit être adapté.  
Lors d'une réunion, les 4 experts, assistés par une autre personne minimum, confrontent leurs traductions. Lorsque des désaccords apparaissent le mot ou la structure de la phrase qui constitue le désaccord est discuté et arbitré.
- (2) Le questionnaire est proposé à des patients (10 à 30) [88-90], représentatifs de la population à laquelle est destiné le questionnaire.  
Chacun des patients va répondre au questionnaire et préciser les difficultés rencontrées : la difficulté de la question, la compréhension, la confusion, les mots ou les phrases perturbantes.
- (3) Suivant ce test, le questionnaire est ensuite réévalué et modifié.

Si la méthode de la traduction-rétro traduction est le plus souvent employée, et fortement conseillée par des instances telles que la FDA ou l'EORTC [90-93], la confrontation des traductions indépendantes n'en n'est pas moins valide, aucune étude n'a montré la supériorité de telle ou telle méthode [88, 89].

Certains auteurs ont même plutôt mis en avant la faiblesse de la traduction-rétro traduction, puisqu'elle serait favorable au mot à mot et ne permettait donc pas une conversation fluide, naturelle et adaptée [94].

## 6.2. La validation

L'instrument nouvellement adapté à la culture d'un pays doit faire l'objet d'une validation psychométrique (Chapitre 2, §5).