

**UNIVERSITE RENE DESCARTES  
PARIS V**

**DEA d'ETHIQUE MEDICALE ET BIOLOGIQUE  
Directeur Pr. C. HERVE  
Promotion 2001-2002**

**UTILISATION DE LA QUALITE DE VIE  
DANS LES ESSAIS DE PHASE III EN CANCEROLOGIE :  
QUALITE METHODOLOGIQUE ET ETHIQUE  
Analyse des essais de phase III publiés de 1999 à 2001 dans 10 revues internationales.**

**Par TUECH Jean-Jacques**

**Directeur de mémoire : Pr SCHRAUB (Strasbourg)**

## **Table des matières :**

<b>Introduction</b>	<b>p 3</b>
<b>Matériels et méthodes</b>	<b>p 5</b>
<b>Résultats</b>	<b>p 8</b>
<b>Tableaux et figures</b>	<b>p 14</b>
<b>Discussion</b>	<b>p 26</b>
<b>Qualité méthodologique</b>	<b>p 31</b>
<b>Qualité éthique</b>	<b>p 34</b>
<b>Qualité de vie</b>	<b>p 38</b>
<b>Conclusion</b>	<b>p 49</b>
<b>Références</b>	<b>p 50</b>

## **Annexes**

**Annexe 1 : Grille de lecture méthodologique selon Jadad (9).**

**Annexe 2 : Grille de lecture éthique selon Berdeu (11)**

**Annexe 3 : grille d'évaluation de Ruiz-Canela (12)**

**Annexe 4 : grille d'évaluation de Gill et Feinstein (13)**

**Annexe 5 : CONSORT statement**

**Annexe 6 : Recommandations éthiques aux auteurs**

**Annexe 7 : Variation de la QOL :schéma selon Calman (79)**

## **Introduction :**

En cancérologie, les tumeurs sont caractérisées par leur diversité, leur histoire naturelle et leur grande hétérogénéité et ce, au sein d'un même groupe de pathologie. Dans un domaine aussi vaste, l'amélioration des connaissances passe par la découverte de nouvelles techniques et de traitements sur la base de données homogènes et irréfutables et ce au-delà de l'observation clinique. Avec Claude Bernard les sciences médicales ont réalisé un bond en avant grâce à l'expérimentation rigoureuse et objective, se basant sur des observations individuelles ou sur un nombre réduit de sujets. Les essais cliniques représentent quant à eux une nouvelle étape dans la progression des connaissances dans le sens où ils proposent une médecine basée sur des faits prouvés. Cochrane avec son concept de l'« evidence based medicine » (EBM) a été le premier à démontrer la nécessité de baser les pratiques sur des données prouvées. Il propose une classification des essais thérapeutiques basée sur la reconnaissance de la valeur hiérarchique et inégale des preuves scientifiques.

L'essai randomisé avec ses groupes comparatifs, obtenus par tirage au sort, permet de faire progresser la connaissance en apportant des arguments basés sur des preuves validées scientifiquement. Il représente l'outil le plus important hiérarchiquement pour juger de l'utilité d'un traitement car il offre trois avantages majeurs : il supprime le biais de sélection des patients ; la randomisation équilibre dans les deux groupes toutes les covariables connues et inconnues ; la randomisation garantit la validité des comparaisons statistiques entre deux groupes puisque le facteur randomisé n'est pas corrélé aux caractéristiques des groupes. Cependant les services rendus par ces outils d'évaluation doivent être considérés avec un esprit critique. Il arrive en effet que la stratégie de l'essai soit mal élaborée parce que la situation choisie est non pertinente cliniquement ou que l'objectif principal est mal adapté à la situation clinique en question, donnant lieu à des essais dont l'utilité est discutable. Mais les principaux reproches faits aux essais contrôlés randomisés (ECR) concernent leur qualité méthodologique et éthique ; les ECR doivent répondre à des exigences cliniques et scientifiques qui sont sous-tendues par des réglementations strictes visant en premier lieu la protection du malade, et dans un second temps l'apport d'un réel bénéfice pour la société.

En cancérologie l'évaluation des thérapeutiques se faisant essentiellement à l'aide de la durée de survie-globale ou sans rechute et de la diminution du volume tumoral (réponse tumorale) corrélées à la l'amélioration des symptômes. Les cancérologues étaient depuis longtemps conscients de l'importance de l'état général du patient ; en 1947 est mis au point le premier index permettant de quantifier le retentissement de la maladie cancéreuse sur

l'état fonctionnel et l'autonomie des patients (1). La notion de qualité de vie (QDV) apparaît en 1970 et en 1996, la société Américaine d'oncologie médicale (ASCO) considérait les mesures de survie et de QDV prioritaires par rapport à l'évaluation de la réponse tumorale (2)

La qualité de vie est par essence un concept subjectif et multidimensionnel ; il apparaît donc très difficile d'en donner une définition standard et unique. En 1993, l'Organisation Mondiale pour la Santé (OMS) a proposé une définition, qui met l'accent sur la notion de perception : « La qualité de vie est définie comme 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 » (3)

Certains auteurs comme Osoba et al (4), considèrent que l'étude de la QDV devrait être incluse pour tous les essais de phase III en cancérologie; d'autres auteurs (5,6) ont une approche plus pragmatique et recommande l'inclusion de l'étude de la QOL pour les essais où la QDV est le critère de jugement principal, comme critère de jugement secondaire pour les essais où la différence attendue en terme de guérison ne devrait pas être significative, et les essais où la différence en terme de survie devrait être importante mais au prix d'une toxicité importante.

Le but de cette étude était d'évaluer la qualité méthodologique et éthique ainsi que la qualité de l'usage et du rapport de la QDV au sein des essais randomisés de phase III en cancérologie.

## **Matériels et méthodes :**

Nous avons identifiés les essais randomisés de phase III traitant de cancérologie publiés entre le 01/01/1999 et le 31/12/2001 dans quatre périodiques médicaux hebdomadaires internationaux de langue anglaise [ New England Journal of Medicine (NEJM), The Lancet, The British Medical Journal (BMJ), The Journal of the American Medical Association (JAMA)] et six périodiques d'oncologie internationaux de langue anglaise [ The British Journal of Cancer, The Journal of Clinical Oncology (JCO), Lung Cancer, Annals of Oncology, The European Journal of Cancer, Clinical Cancer Research]

Les quatre périodiques médicaux ont été choisis en raison de la qualité des essais qu'ils publient et qu'en pratique l'immense majorité des essais décisifs paraît dans une de ces quatre revues ; les six revues spécialisées en oncologie ont été choisies car elles sont l'organe de publications des plus importantes sociétés d'oncologie ( ).

Les essais de phase III étaient définis comme des essais comparatifs avec attribution aléatoire des traitements, le traitement en cours d'évaluation étant comparé au traitement de référence ou à un placebo.

Ont été exclus de cette étude les essais publiés sous forme de lettre, de résumé ou de communication courte, les essais non thérapeutiques, ainsi que les essais randomisés de phase II..

L'identification des essais a été réalisée par une recherche manuelle, compte tenu des réserves, émises sur le manque de fiabilité de la recherche effectuée sur les essais contrôlés randomisés par interrogation de la base de données Medline (7,8).

Les articles rapportant des ECR de phase III ont été photocopiés sans que ces copies aient été rendues anonymes (9).

Le nom du journal, l'année de publication, l'année de début de l'essai, le type de cancer étudié, le nombre de patients inclus, le pays d'origine, le résultat positif ou négatif de l'essai, l'aspect uni- ou multicentrique des inclusions, l'utilisation ou non d'un questionnaire de QDV et son type, ont été recueillis de façon standardisée.

Était considéré comme négatif un essai ne permettant pas le rejet de l'hypothèse nulle, c'est à dire un essai ne permettant pas de montrer que les traitements étaient statistiquement différents.

Parmi cette base de données, ont été sélectionnés les essais étudiant des patients porteurs de cancer à un stade avancé, dans ce type d'essais la QDV pourrait devenir le critère principal d'évaluation des traitements particulièrement quand ceux-ci sont lourds (10). Les patients étudiés ont été considérés comme atteints d'un cancer à un stade avancé lorsque la survie médiane des groupes étudiés était inférieure à 24 mois.

Les articles ainsi sélectionnés ont été évalués sur le plan de la qualité méthodologique, de la qualité éthique et de la qualité de l'utilisation de critères de QDV.

**La qualité méthodologique** a été évaluée par la grille de lecture proposée par Jadad et al (9) en 1996 (annexe 1), cette grille comporte 11 items et le score maximal est de 13. Les 2 premiers items (randomisation, double aveugle) pouvant obtenir un score de 0 à 2, alors que les autres items ne peuvent obtenir un score que de 0 ou 1. Cette grille peut également être utilisée avec seulement les 3 premiers items (score maximum = 5) ou les 6 premiers items (score maximum = 8). La qualité méthodologique d'un ECR était reconnue comme excellente lorsque l'ECR obtenait plus de 4 points à la grille de 6 items ou plus de 2 points à la grille à 3 items. La qualité méthodologique d'un ECR était reconnue comme médiocre lorsque l'ECR obtenait 4 points (ou moins) à la grille de 6 items ou 2 points (ou moins) à la grille à 3 items.

**La qualité éthique** a été évaluée par la grille d'évaluation de la qualité éthique des ECR constituée de 10 items (annexe 2) proposée par Berdeu et al (11), chaque item rapporte 1 point: le score global est obtenu en divisant la somme des scores par le score maximum attribuable. Ce score global est exprimé sous forme d'un nombre décimal compris entre 0 et 1.

Nous avons ajouté à cette grille des items empruntés à la grille (annexe 3) proposée par Ruiz-Canela et al (12)

1-a : Type de consentement. : oral avec témoin ou sans témoin , écrit , non mentionné

1 -b : Information sur l'obtention du consentement. 1)oui 2) non

Ainsi que 5 autres items : Paiement des patients pour la recherche ?

Paiement des chercheurs ?

Financement public ?

Financement privé ?

épidémiologiste / statisticien comme auteurs ?

**Qualité de vie :** La présence ou l'absence d'une étude de la QDV été notée , dans le cas ou l'ECR comportait une étude de la variation de la QDV, la qualité de l'utilisation des critères de QDV était évaluée par la grille d'évaluation de Gill et Feinstein (13) (annexe 4).

Cette grille était complétée par quatre questions :

- 1) un instrument de mesure de la qualité de vie était-il utilisé ? la réponse était positive si les auteurs précisent quel instrument ils avaient utilisé (une référence bibliographique n'était pas considérée comme nécessaire) ; Si un instrument spécifique a l'étude a été utilisé, les auteurs devaient l'avoir reproduit dans l'article.
- 2) Quel type d'instrument était utilisé (générique, spécifique d'une pathologie, spécifique à l'étude) Nous avons utilisé la définition de Patrick et Deyo (14) : un instrument est considéré comme générique si il est applicable à une grande variété de pathologies dont la sévérité et le traitement sont différents et à des groupes et sous-sous groupes de population variables. Un instrument est considéré comme spécifique d'une pathologie si il n'est applicable que pour une pathologie donnée dans un groupe de population donné. Nous avons considéré qu'un instrument spécifique à l'étude était un instrument développé par les auteurs pour leur étude et de ce fait non validé précédemment.
- 3) L'instrument utilisé était il validé ?
- 4) L'instrument utilisé était il adapté à l'hypothèse de l'étude ?

### **Analyse statistique :**

La comparaison des variables qualitatives a été réalisée à l'aide du test du Khi<sup>2</sup> ou en cas de petit effectif à l'aide du test exact de Fisher

La comparaison des valeurs quantitatives a été réalisées par un test t-Student en cas de grand effectif, ou par un test non paramétrique de Mann-Whitney en cas de petit effectif.

La comparaison multiple des valeurs quantitatives a été réalisée par une analyse de variance ANOVA. Au sein de l'ANOVA, les comparaisons de variance deux à deux ont été réalisées par un test F. La recherche d'une corrélation entre deux variable quantitatives a été réalisée par une régression linéaire exprimée par sa droite de régression et par son coefficient de corrélation R<sup>2</sup>.

Une différence a été jugée significative pour une valeur de  $p < 0,05$ .

Toutes ces analyses ont été réalisées à l'aide du logiciel Statview version 5.0

## Résultats :

Entre le 01/01/1999 et le 31/12/2001, seulement deux ECR traitant de cancérologie ont été publiés dans le BMJ, ces deux ECR ne concernant pas des patients porteurs de cancer à un stade avancé, ce journal a donc été exclu de la présente étude, il en était de même pour le JAMA dans lequel un seul ECR traitant de cancérologie a été publié en 1999.

Durant la période étudiée, 604 ECR ont été publiés dans le NEJM (n=235) et le Lancet (n=369), le nombre d'ECR traitant de cancérologie était respectivement de 22 et 21 (tableau 1). Le nombre d'ECR publiés dans les six périodiques d'oncologie était de 216. Un ECR publié dans JCO en 1999 a été exclu de l'étude car il s'agissait d'un ECR faisant appel à des drogues cytotoxiques dans le traitement de maladies rhumatismales (15).

Pour ces 259 ECR, 170.076 patients ont été inclus. Le nombre d'ECR publiés par année et par périodique ont été consignés sur les tableaux 1 .

Ces ECR concernaient les tumeurs du sein dans 26.3% (n=68), les tumeurs pulmonaires dans 14.6% (n=38), les tumeurs urinaires dans 10% (n=26), les effets secondaires des tumeurs ou des traitements dans 8.9% (n=23), les tumeur colorectales dans 8.5% (n=22), les tumeurs hématologiques dans 8.5% (n=22), les tumeurs gynécologiques dans 7.3% (n=19), les tumeurs gastriques dans 4.2% (n=11), les tumeurs cutanées dans 2.7% (n=7), les tumeurs neurologiques dans 2.3% (n=6), les tumeurs primitives hépatiques dans 2% (n=5), les tumeurs de la tête et du cou dans 2% (n=5), les tumeurs de l'œsophage dans 1.5% (n=4), les sarcomes dans 0.8% (n=2) et les carcinomes péritonéales dans 0.4% (n=1).

Les ECR étaient originaire d'Europe dans 55.2% (n=143), d'Amérique du nord dans 28.2% (n=73), de plusieurs continents (internationaux) dans 9.6% (n=25), d'Asie dans 3.9% (n=10), d'Australie et Nouvelle-Zélande dans 2.7% (n=7) et du Qatar dans 0.4% (n=1).

Les ECR étaient multicentriques dans 85% (n=220) , unicentriques dans 15% (n=39), les résultats étaient positifs dans 75% (n=194) et négatifs dans 25% (n=65).

Un questionnaire de QDV était utilisé dans 72 ECR (27.8%), il s'agissait :

-32 fois (43%) de l'EORTC Quality-of-Life Questionnaire C30 (EORTC QLQC30) (16)

-9 fois (12.5%) du Functional Assessment of Cancer Therapy (FACT) (17)

-8 fois (9.7%) du Rotterdam Symptom Checklist (RSCL) (18)

-4 fois (5.5%) d'un questionnaire spécifique

-4 fois (5.5%) du Functional Living Index Cancer (FLIC) (19,20)

-3 fois (4.1%) du Lung Cancer Symptom Scale (21)

-1 fois (1.4%) le SF 36 (22)

-1 fois (1.4%) le Unisquale (23),

-1 fois (1.4%) le Linear Analogue Self-Assessment (LASA) (24) associé au Spitzer QOL Index (25)

-1 fois (1.4%) le LASA (26),

-1 fois (1.4%) le QOL assessment of cancer patients receiving chemotherapy (27)

-1 fois (1.4%) le Perceived Adjustment to Chronic Illness Scale (PACIS) (28)

-1 fois (1.4%) le Eight-item Patient benefit questionnaire (PBQ) (29)

-1 fois (1.4%) le QLQC 30 associé au RSCL

-1 fois (1.4%) le RSCL associé au au Body Image Scale (non encore publié) (P Hopwood. Cancer research campaign Medecine Group, Stanley House, Christie hospital, Manchester UK.)

Dans 5 cas le type de questionnaire utilisé n'était pas précisé dans la publication, le rapport sur la QDV devant faire l'objet d'une publication spécifique .

La proportion d'ECR comportant une étude de la QDV en fonction de l'année de début de chaque ECR a été représentée sur la figure 1. Cette proportion était significativement plus importante lorsque les ECR avaient débuté entre 1995 et 1999 (32/63) que lorsqu'ils avaient débuté entre 1990 et 1994 (28/110),  $\text{Khi}^2$  à 11.29 ( $p=0.00078$ ). Pour 26 ECR (10%), la date de début de l'essai n'était pas mentionnée.

Parmi ces 259 ECR, 84 étudiaient des patients porteurs de cancer à un stade avancé, la survie médiane des groupes étudiés était de 10.35 mois (extrême : 1,6 - 24) ; la médiane de la différence de survie entre les groupes d'un même ECR était de 1,2 mois (extrême : 0 – 14,5).

Le score global de qualité méthodologique (score de Jadad) variait entre 7 et 13 , la moyenne des scores globaux était de  $9,86 \pm 1,117$ . L'analyse critère par critère a été détaillée dans le tableau 2. Le détail des scores obtenus lorsque était utilisé le score à 3,6 ou 11 items a été rapporté sur le tableau 3. Le tableau 4 récapitulait le nombre d'ECR en fonction des scores obtenus .

Pour 26 ECR la méthodologie était insuffisante : ECR ayant obtenu un score global (11 items) inférieur ou égal à 9 et un score « 6 items » inférieur ou égal à 4 et un score « 3 items » inférieur ou égal à 2.

Deux articles pouvaient répondre à l'item n° 2 (double aveugle) abaissant le score maximal possible à 3 pour le score à 3 items, à 6 pour le score à 6 items et à 11 pour le score à 11 items.

Le score global de la qualité éthique variait de 0.11 à 0.66 (tableau 5). La moyenne des scores globaux des 84 ECR était de  $0.366 \pm 0.124$ . L'analyse critère par critère (tableau 6) montrait que les critères les mieux respectés (plus de 70%) étaient les critères 1,2,4 et 7.

Le consentement était obtenu dans 91.6% mais le type de consentement des patients était : non disponible 29 fois (34.5%), écrit 49 fois dont une fois en présence d'un témoin (58.3%), oral avec témoin 3 fois (3.6%), oral sans témoin 2 fois (2.4%), écrit ou oral avec témoin 1 fois (1.2%). Les chercheurs avaient été payés pour leur recherche dans 2.4% (n=2) des ECR. L'ECR était soutenu par un financement public dans 29.7% (n=25) et par un financement privé (industrie pharmaceutique) dans 34.5% (n=29). Un épidémiologiste ou un statisticien avait pris part à l'étude (auteur ou remerciements) dans 46.4% des ECR (n=29).

La moyenne des scores éthiques pour les articles de qualité méthodologique insuffisante (n=26) était de  $0.33 \pm 0.112$  (extrême : 0.11 – 0.66), elle était de  $0.38 \pm 0.127$  (extrême : 0.11 – 0.66) pour les ECR de qualité méthodologique élevée (n = 58). Il n'était pas mis en évidence de différence significative (test de Mann-Whitney) entre ces 2 moyennes (p = 0.056).

L'analyse de variance ANOVA ne mettait pas en évidence d'amélioration de la qualité méthodologique (p=0.62) et de la qualité éthique (p=0.29) en fonction de l'année de démarrage de l'essai. De même il n'était pas mis en évidence de différence de la qualité méthodologique (p=0.23) et de la qualité éthique (p=0.26) en fonction des journaux dans lesquels les ECR étaient publiés.

Il n'existait pas de corrélation linéaire (figure 2) entre le score méthodologique (Jadad) et le nombre de patients inclus dans les ECR ( $R^2=0.003$  ,  $p=0.61$ ).

Par contre il existait une corrélation linéaire entre d'une part (figure 3) le score éthique et le nombre de patients inclus dans les ECR ( $R^2=0.078$  ,  $p=0.01$ ), et d'autre part (figure 4) entre le score éthique et le score méthodologique ( $R^2=0.087$  ,  $p=0.006$ ).

Le recueil du consentement éclairé n'était pas obtenu pour 2 ECR de qualité méthodologique insuffisante ( $n=26$ ) et pour 5 ECR de qualité méthodologique élevée ( $n = 58$ ), il n'était pas mis en évidence de différence significative concernant ce critère :  $p = 0.62$  (test exact de Fisher).

La soumission à un comité d'éthique n'était pas mentionnée pour 9 ECR de qualité méthodologique insuffisante ( $n=26$ ) et pour 12 ECR de qualité méthodologique élevée ( $n = 58$ ), il n'était pas mis en évidence de différence significative concernant ce critère  $\text{Khi}^2$  à 1.83 ( $p=0.17$ ).

Le nombre d'ECR comportant comme critère d'évaluation la QDV était de 37 (44 %), 68 ECR (81%) comportaient une échelle d'évaluation des effets toxiques du traitement. Dans un ECR, la QDV était défini comme un « end-point » dans le chapitre matériel et méthode mais les auteurs ne décrivaient l'utilisation d'aucun instrument de mesure aussi bien au niveau du chapitre matériel et méthode qu'au niveau du chapitre résultat ; dans un autre ECR utilisant le Lung Cancer Symptom Scale, les auteurs ne rapportaient aucun résultat concernant la QDV. Pour 3 ECR, le rapport de l'étude de la QDV était partiel car devant être le sujet d'une autre publication centrée sur la QDV ; Pour 2 autres ECR , le rapport de la QDV était absent pour les mêmes raisons.

Dans 36 cas le type d'instrument d'évaluation de la QDV utilisé était précisé, il s'agissait 18 fois du EORTC-QLQC30 (16), 5 fois du FACT (17), 3 fois du RSCL(18), 3 fois du Lung Cancer Symptom Scale (21), une fois du QLQC30 associé au RSCL, une fois du RSCL associé au Body Image Scale (non encore publié) (P Hopwood. Cancer research campaign Medecine Group, Stanley House, Christie hospital, Manchester UK.), une fois de l'uniscale (23), une fois du LASA (24) associé au Spitzer QOL Index (25) et une fois du LASA (26).

Les instruments de mesure de la QDV étaient sept fois spécifiques d'une pathologie et 29 fois génériques, associés neuf fois à des modules. Dans tous les cas, ces instruments étaient validés et adaptés à l'étude dans 97% des cas. L'analyse critère par critère évaluée par la grille d'évaluation de Gill et Feinstein (13) a été détaillée sur le tableau 7.

L'existence d'une analyse de la QDV n'était dans aucun des 37 ECR mentionnée au niveau du titre, mentionnée dans l'abstract 1 fois (48.6%) et 7 fois (19%) au niveau de la conclusion de l'abstract.

Dans tous les ECR (n=33), les patients complétaient les questionnaires avant le début du traitement (avant ou après la randomisation) puis à un rythme variable. Dans 16 ECR (48.5%) le moment du remplissage par rapport au traitement (avant, pendant, après la cure) n'était pas précisé ; les patients complétaient dans 7 ECR (21.2%) le questionnaire avant le début d'une nouvelle cure, dans 1 ECR (3%) les patients complétaient le questionnaire pendant la cure et dans 9 ECR (27.3%) les patients complétaient le questionnaire après la cure.

Le taux d'observance vis à vis du remplissage des instruments de mesure de la QOL était précisé dans 22 ECR (66.6%), le pourcentage médian était de 81% (extrême : 26 –100). Le problème des « missing data » était évoqué dans le chapitre matériel – méthode et résultats dans 3 cas (9%).

Dans les 33 ECR, le QDV était améliorée 10 fois ; 25 fois il n'y avait pas de différence significative entre les groupes en fin de traitement. On notait 8 fois une différence significative entre les groupes en fin de traitement, soit un groupe présenté une amélioration de sa QDV (n=5), soit un groupe présentait une détérioration de sa QDV (n=3).

L'analyse du chapitre discussion des études publiées montrait dans 9 cas (27.3%) l'absence de discussion concernant la QDV, dans 13 cas la discussion reprenait les résultats en précisant 9 fois (27.3%) qu'il n'y avait pas de différence de QDV entre les groupes et que le ou les traitements ne compromettaient pas la QDV et en précisant dans 4 cas (12.1%) que l'un des traitements avait amélioré la QDV. Dans une étude (3%) les auteurs insistaient (discussion) sur l'amélioration de la QDV dans un des bras alors qu'il n'y avait pas de différence significative exposée au niveau du chapitre résultats.

Au niveau du chapitre discussion de 10 ECR (30.3%), les auteurs exposaient les limites de leur travail quant-à l'évaluation de la QDV ( critique de la méthode, de l'instrument choisi, du rythme de remplissage de questionnaire)

-instrument de mesure non spécifique et questionnaire complété avant les traitements (n=1)

-Biais lié à la différence de toxicité de 2 traitements : l'un ayant une toxicité rapide, l'autre ayant une toxicité cumulative (n=1)

-Biais de remplissage des questionnaires (n=8) responsable d'une surévaluation de la QDV, les malades dont la QDV est altérée ne sont plus capables ou ne souhaitent pas remplir un questionnaire (diminution de la compliance de remplissage plus l'essai avance).

<b>Périodique</b>	<b>Année de publication</b>	<b>ECR (n)</b>	<b>ECR de cancérologie (n)</b>	<b>ECR de cancérologie étudiant la QDV (n)</b>
NEJM	1999	78	11	0
	2000	76	5	1
	2001	81	6	0
Lancet	1999	157	8	1
	2000	119	8	2
	2001	93	5	2
Eur J Can	1999		4	1
	2000		6	4
	2001		10	4
JCO	1999	36	35	13
	2000		51	16
	2001		49	18
Lung Cancer	1999		3	1
	2000		1	1
	2001		4	1
Clin Can Res	1999		1	1
	2000		2	0
	2001		0	0
Ann Oncol	1999		6	1
	2000		10	2
	2001		11	1
Br J Can	1999		6	1
	2000		6	1
	2001		11	0
			<b>259</b>	<b>72</b>

**Tableau 1:** nombre d'ECR, d'ECR de cancérologie et nombre d'ECR comprenant une étude de la qualité de vie (QDV) en fonction du périodique et de l'année de publication.

items	score		
	0	1	2
1.L'étude était elle décrite comme étant randomisée	0	37 (44)	47 (56)
2. L'étude était elle décrite comme étant en double aveugle.	82 (97.6)	0	2 (2.4)
3. existait il une description des retraits après randomisation	18 (21.4)	66 (78.6)	
4. les objectifs de l'étude étaient-ils clairement définis ?	4 (4.8)	80 (95.2)	
5. les méthodes de mesure des résultats étaient-elles clairement définies ?	1 (1.2)	83 (98.8)	
6. les critères d'inclusion et d'exclusion étaient-ils clairement définis ?	1 (1.2)	83 (98.8)	
7. la taille de l'échantillon nécessaire était-elle justifiée ?	13 (15.5)	71 (84.5)	
8. les interventions étaient-elles bien décrites ?	0	84 (100)	
9. existait-il au moins un groupe contrôle (comparatif) ?	0	84 (100)	
10. La méthode pour évaluer les effets indésirables était-elle bien décrite ?	20 (23.4)	64 (76.2)	
11. la méthode d'analyse statistique était-elle bien décrite ?	3 (3.6)	81 (96.4)	

**Tableau 2** : grille méthodologique de Jadad, analyse critère par critère des 84 ECR étudiants des patients porteurs de cancer à un stade avancé. Les chiffres entre parenthèses sont des pourcentages.

ECR	3 items	6 items	11 items	ECR	3 items	6 items	11 items
1	3	6	9	43	3	6	11
2	<b>3</b>	<b>6</b>	<b>10</b>	44	<b>3</b>	<b>6</b>	<b>10</b>
3	2	5	10	45	3	6	11
4	<b>3</b>	<b>5</b>	<b>10</b>	46	<b>2</b>	<b>4</b>	<b>9</b>
5	2	5	10	47	1	4	8
6	<b>3</b>	<b>6</b>	<b>11</b>	48	<b>2</b>	<b>5</b>	<b>10</b>
7	3	6	9	49	2	5	10
8	<b>3</b>	<b>6</b>	<b>11</b>	50	<b>2</b>	<b>5</b>	<b>10</b>
9	1	4	9	51	2	5	10
10	<b>2</b>	<b>5</b>	<b>9</b>	52	<b>2</b>	<b>5</b>	<b>10</b>
11	2	4	9	53	2	5	10
12	<b>2</b>	<b>5</b>	<b>10</b>	54	<b>3</b>	<b>6</b>	<b>11</b>
13	3	6	10	55	3	6	11
14	<b>2</b>	<b>5</b>	<b>8</b>	56	<b>2</b>	<b>5</b>	<b>8</b>
15	1	4	8	57	3	6	9
16	<b>3</b>	<b>6</b>	<b>11</b>	58	<b>2</b>	<b>5</b>	<b>10</b>
17	2	5	10	59	2	5	10
18	<b>3</b>	<b>6</b>	<b>11</b>	60	<b>2</b>	<b>5</b>	<b>10</b>
19	2	5	9	61	2	5	10
20	<b>3</b>	<b>6</b>	<b>10</b>	62	<b>2</b>	<b>5</b>	<b>8</b>
21	3	6	11	63	3	6	10
22	<b>3</b>	<b>6</b>	<b>11</b>	64	<b>3</b>	<b>6</b>	<b>11</b>
23	2	5	10	65	2	5	10
24	<b>3</b>	<b>6</b>	<b>11</b>	66	<b>2</b>	<b>5</b>	<b>9</b>
25	1	4	9	67	3	6	11
26	<b>3</b>	<b>6</b>	<b>11</b>	68	<b>3</b>	<b>6</b>	<b>11</b>
27	1	4	9	69	3	6	11
28	<b>1</b>	<b>4</b>	<b>9</b>	70	<b>2</b>	<b>5</b>	<b>8</b>
29	2	4	7	71	3	6	10
30	<b>3</b>	<b>6</b>	<b>10</b>	72	<b>3</b>	<b>6</b>	<b>9</b>
31	5	8	13	73	2	5	10
32	<b>3</b>	<b>6</b>	<b>11</b>	74	<b>1</b>	<b>4</b>	<b>9</b>
33	2	5	10	75	2	5	10
34	<b>2</b>	<b>5</b>	<b>7</b>	76	<b>2</b>	<b>5</b>	<b>10</b>
35	2	5	10	77	3	6	11
36	<b>2</b>	<b>5</b>	<b>9</b>	78	<b>3</b>	<b>6</b>	<b>11</b>
37	2	5	10	79	3	6	11
38	<b>2</b>	<b>4</b>	<b>9</b>	80	<b>2</b>	<b>5</b>	<b>11</b>
39	2	4	7	81	3	6	11
40	<b>2</b>	<b>5</b>	<b>8</b>	82	<b>2</b>	<b>5</b>	<b>10</b>
41	3	6	11	83	3	6	11
42	<b>3</b>	<b>6</b>	<b>10</b>	84	<b>3</b>	<b>6</b>	<b>11</b>

**Tableau 3:** grille méthodologique de Jadad, score obtenu pour chaque ECR en fonction du nombre d'items pris en compte.

	score	1	2	3	4	5	6	7	8	9	10	11	12	13
<b>3 items</b>		7	39	37		1								
<b>6 items</b>					12	35	36		1					
<b>11 items</b>								3	7	16	32	25		1

**Tableau 4:** tableau récapitulant le nombre d'ECR en fonction du score obtenu pour la grille à 3,6 et 11 items.

**Tableau 5** : (première partie) Score éthique obtenu pour chaque ECR

Critères	1	2	3	4	5	6	7	8	9	10	Score global article
1	oui	oui	non	oui	non		non	oui	non	non	0.44
2	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>0.44</b>
3	oui	oui	oui	oui	non		non	non	non	non	0.44
4	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.44</b>
5	oui	NA	non	oui	non		non	non	non	oui	0.33
6	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
7	oui	oui	non	oui	non		non	non	non	non	0.33
8	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>0.55</b>
9	oui	oui	non	oui	non		non	non	non	non	0.33
10	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>0.55</b>
11	oui	NA	non	oui	non		non	oui	non	non	0.33
12	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
13	NA	NA	non	oui	non		non	non	non	non	0.11
14	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
15	oui	oui	non	oui	non		non	non	non	non	0.33
16	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.44</b>
17	oui	oui	non	oui	non		non	non	non	non	0.33
18	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
19	oui	NA	non	oui	non		non	non	non	non	0.22
20	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>0.55</b>
21	oui	oui	non	oui	non		non	non	non	non	0.33
22	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
23	NA	oui	non	oui	non		non	non	non	oui	0.33
24	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>0.44</b>
25	oui	NA	non	oui	non		non	oui	non	non	0.33
26	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
27	oui	NA	non	oui	non		non	non	non	non	0.22
28	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>0.44</b>
29	oui	oui	non	oui	non		non	non	non	non	0.33
30	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>0.55</b>
31	oui	oui	oui	oui	non	non	non	oui	non	oui	0.60
32	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
33	oui	oui	non	oui	non		non	non	non	non	0.33
34	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
35	oui	NA	non	oui	non		non	non	non	non	0.22
36	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
37	oui	oui	non	oui	non		non	non	non	non	0.33
38	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
39	oui	oui	non	oui	non		non	non	non	non	0.33
40	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
41	oui	oui	non	oui	non		non	oui	non	non	0.44
42	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
43	oui	oui	non	oui	non		non	oui	oui	non	0.55
44	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>0.44</b>
45	oui	oui	non	oui	non		non	oui	oui	oui	0.66
46	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
47	oui	oui	non	oui	non		non	non	non	non	0.33
48	<b>NA</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.11</b>
49	oui	oui	non	oui	non		non	non	non	non	0.33
50	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
51	oui	oui	non	oui	non		non	non	non	non	0.33
52	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>		<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
53	oui	oui	non	oui	non		non	non	non	non	0.33

Critères	1	2	3	4	5	6	7	8	9	10	Score global
<b>Articles</b>											
<b>54</b>	<b>NA</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.11</b>
55	oui	NA	non	oui	non	non	non	non	non	non	0.22
<b>56</b>	<b>NA</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
57	oui	oui	non	oui	non	non	non	non	non	non	0.33
<b>58</b>	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
59	oui	oui	non	oui	non	non	oui	non	non	non	0.44
<b>60</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.55</b>
61	oui	oui	non	oui	non	non	oui	oui	non	non	0.55
<b>62</b>	<b>NA</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.11</b>
63	oui	oui	non	oui	non	non	non	non	non	non	0.33
<b>64</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
65	oui	oui	non	oui	non	non	non	non	non	non	0.33
<b>66</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>oui</b>	<b>oui</b>	<b>oui</b>	<b>0.66</b>
67	oui	oui	non	oui	non	non	oui	oui	non	non	0.55
<b>68</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
69	oui	oui	non	oui	non	non	oui	non	non	non	0.44
<b>70</b>	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
71	oui	oui	non	oui	non	oui	non	non	non	non	0.44
<b>72</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>0.44</b>
73	oui	NA	non	oui	non	non	oui	oui	oui	oui	0.55
<b>74</b>	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.22</b>
75	oui	oui	non	oui	non	non	non	non	non	non	0.33
<b>76</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
77	oui	NA	non	oui	non	non	non	non	non	non	0.22
<b>78</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.55</b>
79	oui	oui	non	oui	non	non	oui	oui	non	non	0.55
<b>80</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>non</b>	<b>0.33</b>
81	NA	NA	non	oui	non	non	oui	oui	oui	oui	0.44
<b>82</b>	<b>oui</b>	<b>NA</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.44</b>
83	oui	oui	non	oui	non	non	non	non	non	non	0.33
<b>84</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>oui</b>	<b>oui</b>	<b>non</b>	<b>non</b>	<b>0.55</b>

**Tableau 5** : (deuxième partie) Score éthique obtenu pour chaque ECR

<b>Critères</b>		
<b>1</b>	Recueil consentement éclairé	Oui : n = 77 (91.6) Non : n = 7 (8.4)
<b>2</b>	Soumission comité éthique	Oui : n = 61 (72.6) Non : n = 23 (27.4)
<b>3</b>	Evaluation au préalable du rapport bénéfice risque individuel et collectif	Oui : n = 2 (2.4) Non : n = 82 (97.6)
<b>4</b>	Respect de la clause d'équivalence a priori	Oui : n = 84 (100%)
<b>5</b>	Taux de refus de consentements	Non précisé : 100%
<b>6</b>	Justification éthique de l'emploi d'un placebo*	Oui : 1 Non : 1
<b>7</b>	Critères d'exclusion abusifs ou non justifiés	Oui : 1 Non : 83
<b>8</b>	Analyses intermédiaires prévues	Oui : 26 (31) Non : 58 (69)
<b>9</b>	Existence règles d'arrêt précoce	Oui : 16 (19) Non : 68 (81)
<b>10</b>	Existence d'un comité indépendant de surveillance de l'essais	Oui : 12 (14.3) Non : 72 (85.7)

**Tableau 6** : Analyse éthique critère par critère selon la grille de Berdeu et al.

Critères		1	2	3	4	5	6	7	8	9 et 10
Type d'instrument	instrument									
spe	LASA	oui	non	non	non	non	non	non	non	non
gen + module	QLQc30	oui	oui	non	non	oui	oui	non	oui	non
gen	QLQc30	oui	oui	oui	oui	oui	oui	oui	oui	non
gen	uniscale	oui	oui	non	oui	non	oui	oui	non	non
gen	RSCL	oui	oui	non	oui	non	non	non	non	non
gen	qlqc30	oui	oui	non	non	non	oui	oui	non	non
gen	qlqc30+RscI	oui	oui	non	non	<u>publ a</u>	<u>part</u>	oui		
gen	QLQc30	oui	oui	non	non	non	oui	oui	non	non
gen+spe	FACT CNS	oui	oui	non	oui	oui	non	non	non	non
gen+spe	Qlqc30	oui	oui	non	non	non	oui	oui	non	non
gen	fact	oui	oui	oui	non	non	oui	non	non	non
gen	qlqc30	oui	oui	non	oui	non	oui	oui	non	non
gen + module	QLQc30	oui	oui	non	oui	non	oui	oui	oui	non
spe	Fact L	oui	oui	non	oui	oui	oui	non	non	non
gen	QLQc30	oui	oui	non	oui	non	oui	oui	oui	non
gen + module	QLQc30	oui	oui	non	non	non	oui	oui	non	non
gen + module	QLQc30	oui	oui	non	oui	non	oui	oui	non	non
gen	QLQc30	oui	oui	non	oui	non	oui	oui	non	non
Gen +module	qlqc30+lc13	oui	oui	non	oui	non	non	oui	non	non
gen	QLQc30	oui	oui	non	non	non	oui	oui	non	non
gen	FACT G	oui	oui	non	non	non	oui	non	non	non
gen + module	QLQc30	oui	oui	non	non	non	oui	oui	non	non
gen	QLQc30	oui	oui	non	non	non	oui	oui	non	non
gen + module	RSCL+BIS	oui	oui	oui	oui	non	oui	oui	oui	non
gen	QLQc30	oui	oui	non	oui	non	oui	oui	non	non
gen	QLQc30	oui	oui	non	oui	non	oui	oui	non	non
spe	Lung cancer symptom scale	oui	oui	non	oui	oui	oui	non	non	non
gen	RSCL	oui	oui	non	oui	non	oui	non	non	non
spe	FACT L	oui	oui	non	non	non	oui	non	non	non
gen	QLQc30	oui	oui	oui	oui	non	non	oui	non	non
gen	RSCL	oui	oui	non	non	non	non	non	non	non
spe	Lung cancer symptom scale	oui	oui	non	non	non	non	non	non	non
gen	spitzer +linear analog scale	oui	oui	non	oui	non	oui	oui	non	non
<b>N=33</b>		<b>33</b> 100%	<b>32</b> 97%	<b>4</b> 12%	<b>18</b> 54.5%	<b>5/32</b> 15.6%	<b>25/32</b> 78%	<b>21/32</b> 65.6%	<b>4/32</b> 12.5%	<b>0/32</b> 0%
type gen spe nom		Validé	adap té?	conc ept qol	domai nes	argum ents choix	score unique	score global	qol/ hqol	ecrit

**Tableau 7:** Analyse critère par critère des ECR étudiant la QOL, évaluation selon la grille de Gill et Feinstein (13).

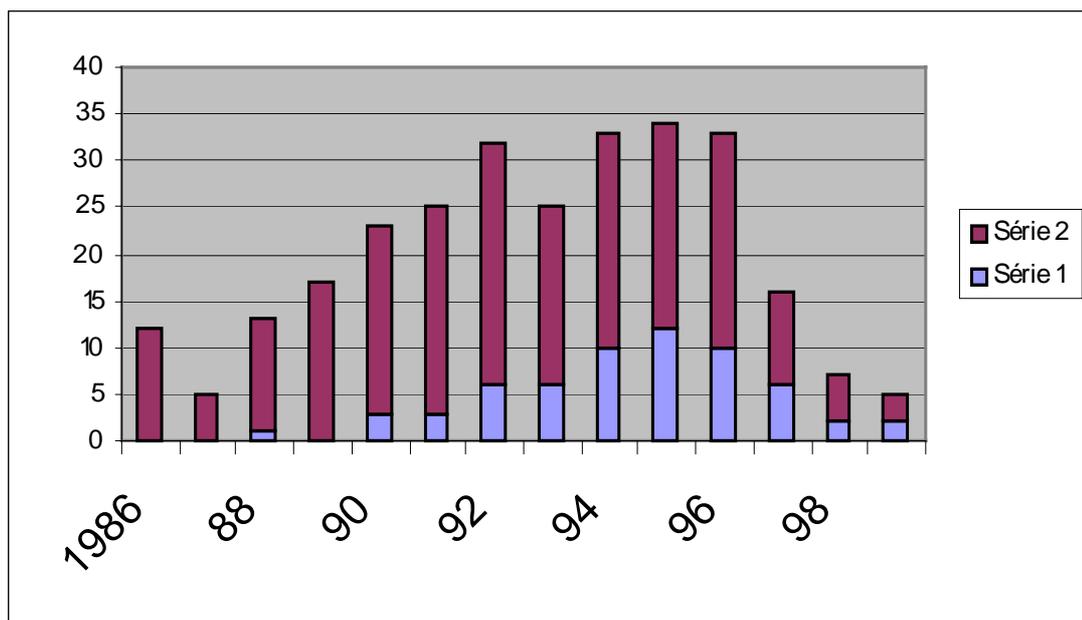
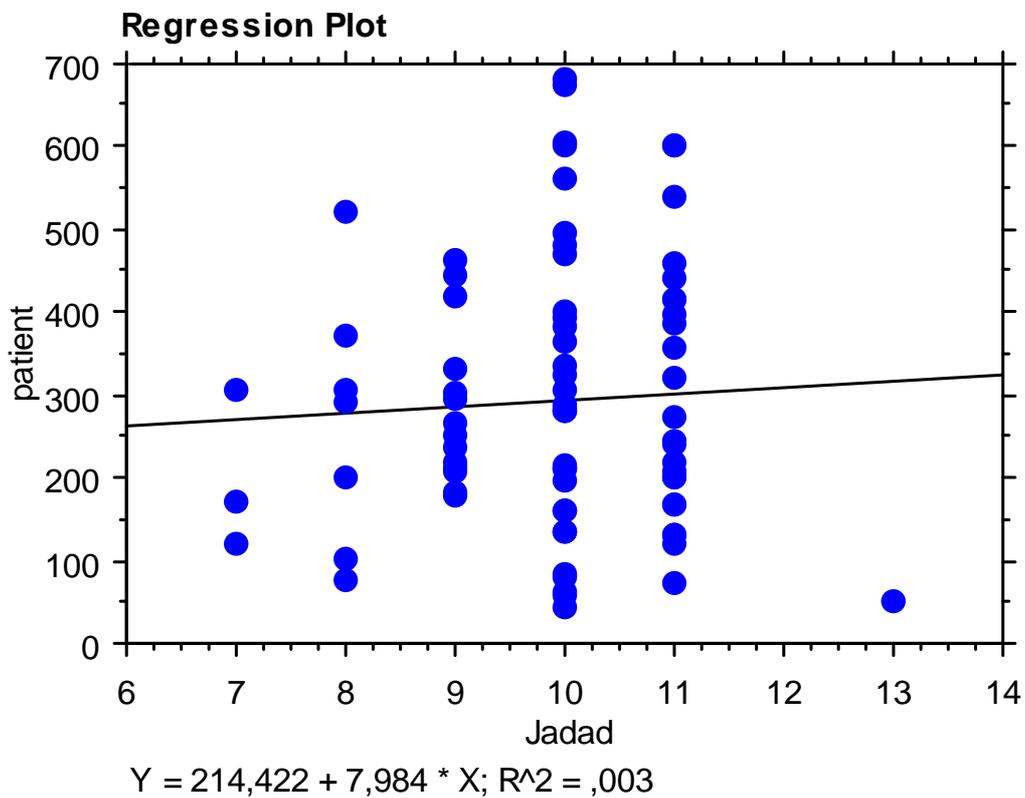
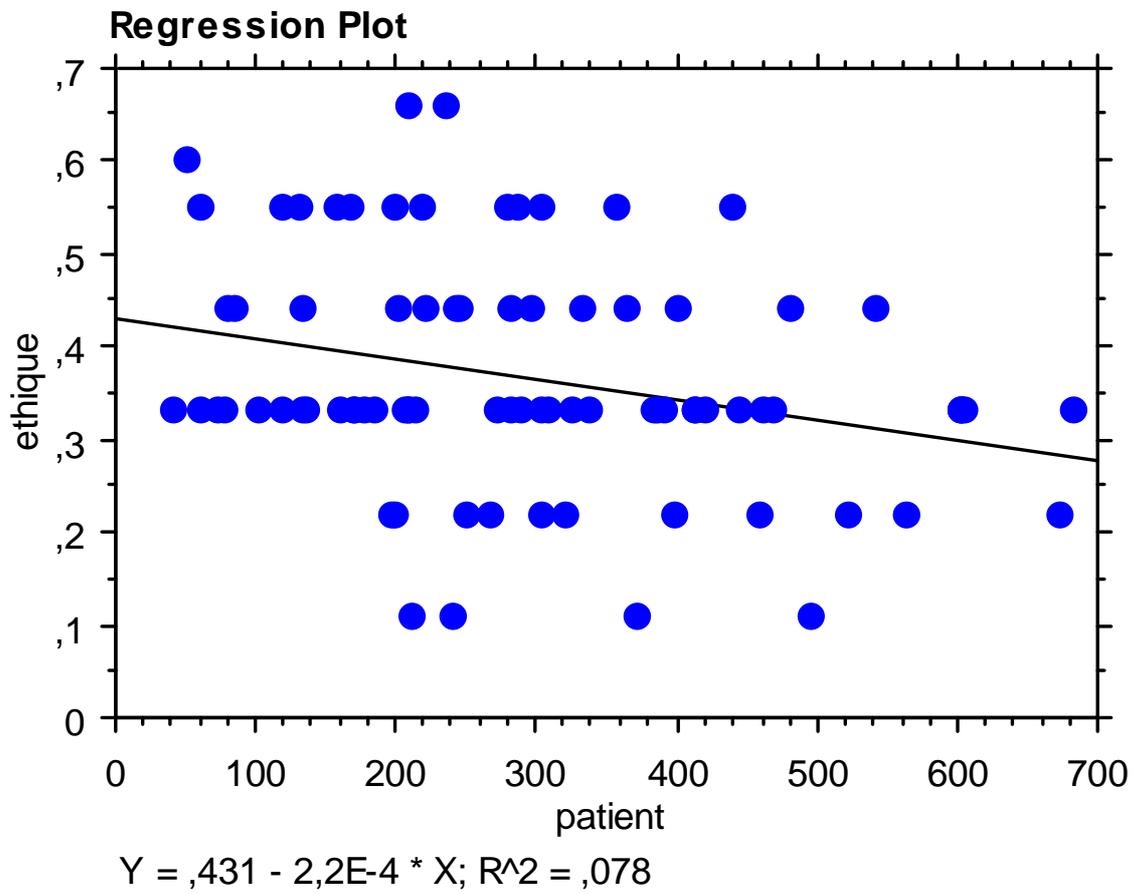


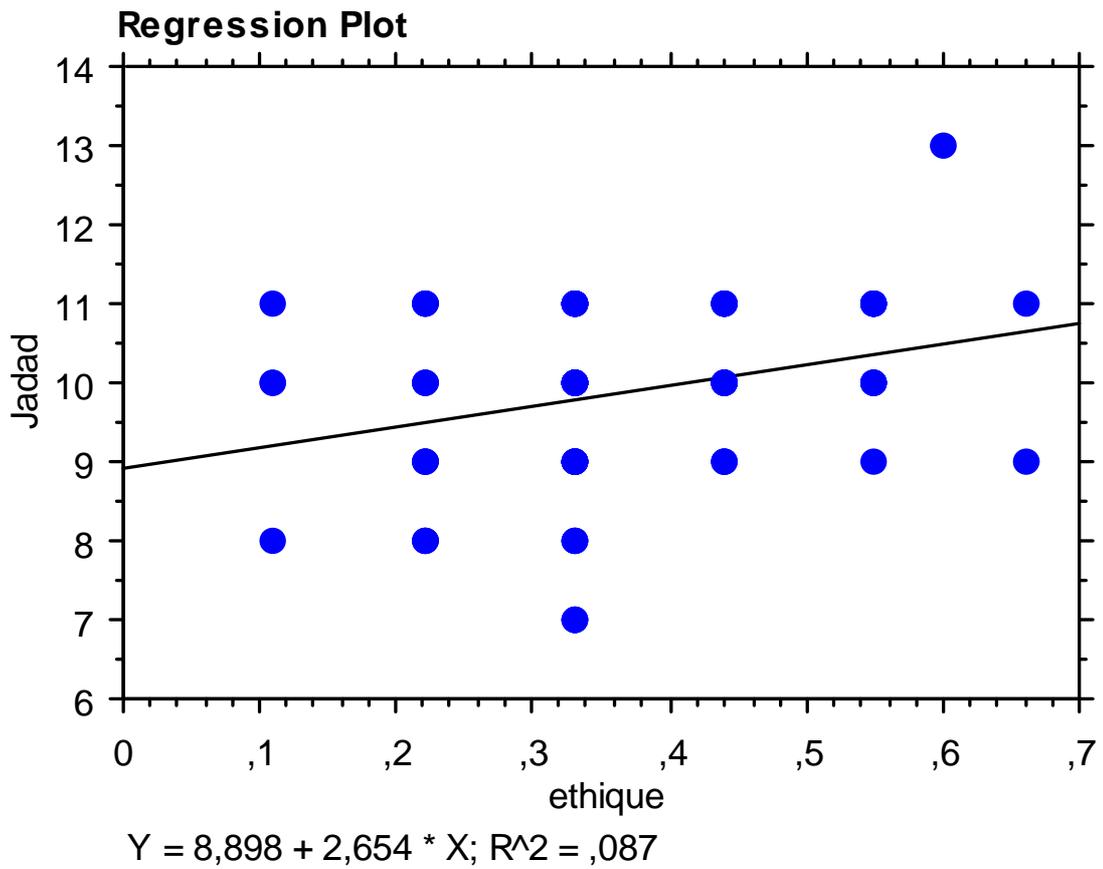
Figure 1 : évolution de la proportion d'ECR avec étude la QDV en fonction de l'année de début des ECR.( série 1 : ECR avec QDV, série 2 : ECR sans QDV)



**Figure 2** : corrélation linéaire entre le score méthodologique (Jadad) et le nombre de patients inclus dans les ECR ( $R^2=0.003$  ,  $p=0.61$ ).



**Figure 3** : corrélation linéaire entre le score éthique et le nombre de patients inclus dans les ECR ( $R^2=0.078$  ,  $p=0.01$ )



**Figure 4** : corrélation linéaire entre le score éthique et le score méthodologique ( $R^2=0.087$  ,  $p=0.006$ ).

## **Discussion :**

La méthode habituelle pour faire circuler l'information est l'écrit. La tradition de l'écrit est très ancienne, elle remonte bien avant l'invention de la presse à imprimer de Gutenberg (1450) ; déjà Moïse l'avait préférée pour transmettre correctement les Dix Commandements. La valeur de l'écrit n'est plus à prouver : elle permet de garder une trace de ce qui a été fait, de ce qui a été dit, de l'histoire des communautés, elle donne une base commune à la connaissance et elle est le moyen par lequel nous partageons les informations qui nous semblent importantes comme les nouvelles du jour ou les dernières découvertes scientifiques (30).

La science nécessite, beaucoup plus que dans d'autres domaines, un vocabulaire précis et clair pour la description des travaux réalisés ainsi que leurs résultats. De plus, pour avancer elle a besoin que soient rapportées toutes les expériences et études conduites et également la manière dont elles ont été menées.

Aujourd'hui la tendance va vers l' « evidence based medicine », ce qui signifie médecine basée sur les preuves, c'est une médecine qui se veut plus scientifique en s'appuyant notamment sur des études qui ont amené la preuve que le traitement A est plus efficace que le traitement B. Nous assistons aujourd'hui à une profusion d'articles rapportant des essais de plus ou moins bonne qualité (méthodologie parfois douteuse, nombre de patients insuffisant,...) avec des résultats plus ou moins divergents, la question se pose alors au clinicien de savoir quelles informations retenir pour la décision thérapeutique. Une des solutions apportées pour résoudre ce problème est la méta-analyse : procédure visant à rassembler tous les articles portant sur un même sujet, de méthodologie similaire afin de former une seule analyse qui atteindra alors une puissance statistique plus élevée (31) .

En France, Le projet " Standards, Options et Recommandations " (SOR) en oncologie a été initié par la FNCLCC (Fédération Nationale des Centres de Lutte Contre le Cancer) en 1993. L'objectif de l'opération SOR est d'améliorer la qualité et l'efficacité des soins aux patients atteints de cancer en fournissant aux praticiens une aide à la décision facilement utilisable à partir d'une revue et d'une analyse critique des données actuelles de la science par un groupe de travail multidisciplinaire. Le projet comprend la définition de recommandations pour la pratique clinique en oncologie, leur diffusion vers les cliniciens et l'évaluation de leur impact sur la pratique clinique. Les SOR ne se substituent pas au jugement du médecin mais apportent une aide à la décision médicale. La méthodologie d'élaboration des SOR repose sur une revue et une analyse critique des données de la littérature scientifique par un groupe d'experts pluridisciplinaire, permettant

de définir, sur la base du niveau de preuve scientifique et du jugement argumenté des experts, des « Standards », des « Options » et des « Recommandations ».

La cancérologie est un domaine où la recherche clinique est omniprésente, utilisant des traitements lourds chez des patients fragilisés tant sur le plan physique que psychique. Les essais thérapeutiques sont très largement répandus dans cette discipline :

- Les essais de phase I : expérimentent l'introduction de nouvelles drogues chez des sujets malades, volontaires, auxquels aucune thérapeutique active ne peut être proposée. Leurs objectifs primaires portent sur la détermination de la dose maximale tolérée et des effets secondaires. Leurs résultats sont souvent tenus secrets en raison des enjeux économiques.
- Les essais de phase II : évaluent sur des petits effectifs homogènes l'efficacité d'un produit et déterminent la posologie optimale pour les essais de phase III sont des essais comparatifs reposant sur l'attribution aléatoire des traitements. Le traitement testé est comparé au traitement de référence dans l'indication thérapeutique étudiée ou à un placebo en l'absence de traitement de référence.

Cette étude a été limitée au champ des ECR de phase III en cancérologie. Ce choix avait pour but la création d'un groupe homogène de documents et de pathologies permettant une analyse standardisée. Les ECR publiés dans 10 revues internationales ont été pris en compte, ce choix restrictif a été guidé par la qualité reconnue des 4 revues médicales et par la représentativité du « monde » de la cancérologie ( X sociétés savantes) pour les 6 revues d'oncologie. Ce choix arbitraire peut avoir introduit un biais responsable d'une surestimation de la qualité des ECR analysés dans l'étude. Toutefois, l'expérience montre que la plupart des ECR en cancérologie sont publiés dans ces revues. Dans notre étude, la majorité des ECR était publiée dans des revues spécifiques (83.4 %) et 52 % (n = 135) était publié dans un seul journal (JCO). Contrairement aux autres disciplines, les cancérologues préfèrent publier leurs ECR dans des revues spécialisées, en 3 ans les 2 revues médicales ont publié 604 ECR dont seulement 43 traités de cancérologie.

Les articles ont été photocopiés sans que les copies aient été rendues anonymes. En effet deux études (9 , 32) ont montré que l'anonymisation ou non des articles ne modifiait pas les scores obtenus lors de leur évaluation méthodologique. Le biais dû à l'effet de masquer ou non l'identité des auteurs a été par ailleurs largement étudié dans la procédure du « peer reviewing », en effet les relecteurs peuvent rejeter injustement un article à cause d'un conflit d'intérêt ou bien accepter ou rejeter un article en raison d'une sensibilité envers l'auteur (respect, amitié, mépris..). Justice et al (33) n'ont pas mis en évidence de différence

significative sur la décision éditoriale dans un essai randomisé réalisé sur 5 journaux de spécialités différentes comparant l'effet de masquer ou non le nom des auteurs. Une autre étude a évalué l'effet de l'anonymisation de l'auteur et du relecteur sur l'acceptation de l'article, aucune différence n'a été mise en évidence entre les relecteurs masqués et les non-masqués. 46% des relecteurs ont pu deviner correctement le nom de l'auteur au travers des références et grâce à leur connaissance du domaine (34).

Dans cette étude, 55.2 % des ECR étaient originaires d'Europe et 28.2% étaient originaire des USA, des proportions semblables avaient été rapportées par Mariani et Marubini (respectivement 50.3 et 37.7%) (35) au sein d'une analyse de 308 essais de phase II publiés antérieurement à 1995. Cette constatation montre que l'Europe a une place prépondérante dans la recherche clinique en oncologie, malgré cela cette notion semble être méconnue face à la puissante organisation qu'est l'ASCO qui organise chaque année un congrès incontournable et dont le périodique (JCO) publie la majorité des ECR. Il serait souhaitable qu'une réflexion s'engage au niveau Européen afin de soutenir, promouvoir et coordonner la recherche clinique. En effet le nombre important de malades à inclure dans les ECR impose la participation de plusieurs centres recruteurs pour chaque ECR, dans cette étude 85% des ECR étaient multicentriques. Le caractère multicentrique ainsi que le nombre élevé de patients inclus impose la gestion des essais par une organisation ou un groupe coopératif ; ceci demande des moyens financiers importants dont les mécènes sont soit l'industrie pharmaceutique soit les financements publics.

Dans cette étude, 75% des ECR étaient positifs (mise en évidence d'une différence significative) et 25 % étaient négatifs. Ceci suggère que les essais positifs sont plus souvent publiés que les essais négatifs, nous renvoyant à la notion de biais de publication. Le premier à avoir soulevé la problématique du biais de publication semble être Diagoras de Melos un poète grec du Vème siècle. Diagoras était un athée et un libre penseur. Sa contribution au biais de publication est liée à une visite dans un temple de l'île de Samothrace en mer Egée. Tous ceux qui avaient survécu à un naufrage ou qui avaient été sauvés de la noyade devaient exposer leurs portraits dans une galerie en guise de remerciements au grand dieu de la mer Neptune. Un croyant essaya de convaincre Diagoras de l'existence des dieux et lui dit « ces portraits ne sont-ils pas la preuve que les dieux interviennent dans les affaires humaines ? », ce à quoi Diagoras répondit « Oui, mais... où sont les portraits de ceux qui ont péri ? ». Cette anecdote avait été reprise par Francis BACON (36). Il est l'auteur d'un ouvrage utopique célébrant la science moderne. Il fait allusion au biais de publication en soulignant que la nature humaine est plus sensible aux

résultats positifs : « l'affirmatif ou l'actif ont plus d'effets que le négatif ou le privatif. De telle manière que quelques réussites contrebalancent un nombre plus élevé d'échecs ou d'absences. ». C'est une description très claire de ce qu'est le biais de publication. Bien que le biais de publication soit très souvent cité, les preuves de son existence sont rares. Un exemple des conséquences du biais de publication a été rapporté en 1993 (37); les antiarythmiques étaient prescrits après infarctus du myocarde afin de prévenir le décès par extra-systoles-ventriculaires (ESV) car il avait été démontré que ces médicaments diminuaient le nombre d'ESV. En 1991, l'essai CAST (38) mettait en évidence une mortalité de 9% avec un traitement contre 4% sans traitement. Ces prescriptions auraient été responsables aux USA d'autant de décès que les guerres de Corée et du Viet-Nam réunies (39). Cette surmortalité aurait pu être évitée car en 1980 un essai non-publié avait déjà observé une augmentation de la mortalité (37) lorsque les patients étaient traités par un antiarythmique.

Easterbrook (40) étudiant le devenir des ECR approuvés par le comité d'éthique d'Oxford (1984-1987) mettait en évidence que 73% des essais avaient été publiés, le nombre d'articles était supérieur lorsque les résultats étaient positifs. La cause de non publication était le résultat négatif de l'ECR dans 21% des cas et le refus des éditeurs dans 9% des cas. Il n'était pas mis en évidence de différence de publication en fonction du mode de financement public ou privé.

Pour Dickersin (41) le refus de publication par l'éditeur variait entre 3.1 et 6.8%. Stern (42) dans une étude similaire mettait en évidence une différence dans les délais de publication en fonction du caractère positif ou négatif de l'ECR ; le délai moyen avant publication était respectivement de 4.82 et 7.99 ans.

Dans cette étude ; 34.5 % (n=29) des ECR étaient financés par les firmes pharmaceutiques dont 22 ECR avaient un résultat positif. Il semblerait que les études financées par des entreprises à but lucratif donnent plus souvent des résultats positifs. Deux raisons ont été avancées pour expliquer cela : la qualité des recherches financées par le privé serait inférieure et par conséquent le bénéfice thérapeutique serait surestimé, une autre hypothèse serait que l'industrie ne finance que les recherches dont elle pense que le résultat sera positif (43). Cette dernière hypothèse est la plus gênante du point de vue éthique : elle consiste à enrôler des patients dans un protocole alors que l'on sait quel va être le résultat.

De nombreux autres biais de publication influencent la représentativité des articles publiés, sans vouloir être exhaustif nous citerons :

- Le biais de la langue de publication : une étude allemande (44) a été réalisée sur des paires d'articles écrit par le même auteur : l'un en anglais, l'autre en allemand. Cette étude a montré que les auteurs publiaient plus souvent dans les revues anglo-saxonnes lorsque les résultats étaient significatifs (35% des articles en allemand sont significatifs contre 62% dans les revues en anglais), en revanche il ne semble pas avoir de différence de qualité entre les articles publiés en allemand et en anglais. L'analyse multivariée mettait en évidence que la seule variable prédictive de publication en langue anglaise était le sens des résultats.

- Le biais de publication multiple : une même étude peut-être publiée plusieurs fois soit en publiant le même article dans différentes revues soit en découpant l'étude en sous-groupes. Cette pratique a pour conséquences de rendre les résultats moins lisibles mais aussi d'influencer sur les résultats des méta-analyses, en effet ce phénomène serait responsables d'une surestimation de l'efficacité d'un médicament de 23% (45)

Rennie (46) au cours d'une méta-analyse a montré que 31 études avaient donné lieu à 75 publications (2 publications: n = 20, 3 publications: n = 10, 5 publications: n = 1); dans 72% des cas aucune mention n'était faite de l'existence de publication antérieure.

Il semble donc que le principal responsable des différents biais de publication soit l'investigateur (autocensure), il faudrait sensibiliser les auteurs aux conséquences pouvant résulter de telles pratiques afin qu'ils adoptent une attitude « fair-play » ou bien réglementer la publication :

- Les auteurs devraient s'engager à publier les résultats lors de la soumission de l'ECR devant le comité de protection des personnes sous peine de se voir interdire la possibilité de réaliser ou de participer à d'autres protocoles.
- Les revues devraient s'engager à publier les essais aux résultats négatifs
- Les comités de protection des personnes devraient avoir la possibilité du suivi des ECR jusqu'à leur publication
- Une étude publiée plusieurs fois ne devrait être prise en compte que comme une seule étude lors d'une épreuve de titres et travaux.

Un questionnaire de QDV était utilisé dans 27.8% (n = 72) des ECR identifiés pour cette étude. On notait une nette prédominance du QLQC30 (43%), le FACT et le RSCL étaient utilisés respectivement dans 12.5 et 9.7%, le FLIC était utilisé dans 5.5%.

Les questionnaires ont été mis au point dans différents pays d'Amérique du nord et d'Europe. Comme nous l'avons vu, la majorité des ECR sont multicentriques et il est nécessaire de disposer de questionnaires internationaux. Ainsi se sont progressivement imposés en Amérique du Nord le FLIC puis le FACT et en Europe le questionnaire mis au point par l'EORTC (QLQC30). Ces questionnaires ont été mis au point pour la plupart en langue anglaise et il est nécessaire que soit faite une adaptation culturelle pour chaque pays du globe ; ceci requiert une méthodologie rigoureuse afin de vérifier que la validité du questionnaire traduit est bien maintenue.

La prépondérance de l'utilisation du QLQC30 dans cette étude pourrait en partie s'expliquer en partie par le nombre important d'essais en provenance d'Europe (55.2%)

Une augmentation continue du pourcentage des ECR évaluant la QDV a été mise en évidence (figure 1) en fonction de l'année de démarrage des essais. La diminution de la proportion d'ECR évaluant la QDV visible au niveau de la figure 1 à partir de 1996 est lié au fait que la majorité des ECR ayant débutés en 1996 et plus récemment sont encore en cours et non publiés. Plusieurs auteurs (47-49) ont souligné l'accroissement de l'utilisation des questionnaires de QDV, leur utilisation était quasiment nulle (0 à 1.5%) en 1980 (47, 48), de 3% en 1995 (47) et de 8.2% en 1997 (48) ; Dans cette étude la proportion d'ECR évaluant la QDV est passée de 25.5 % avant 1995 à 50.8 % après 1995 (année de démarrage des essais et non année de publication). Ces données sont en faveur d'une prise en compte des recommandations de l'ASCO de 1996 (2) rendant la mesure de la QOL prioritaire par rapport à l'évaluation de la réponse tumorale.

La qualité méthodologique est la première exigence éthique à laquelle doivent répondre les ECR (50). L'évaluation de la qualité méthodologique est un art difficile dont le but est de déterminer si l'absence de biais est suffisamment probable pour que l'on puisse considérer son résultat comme une estimation exacte de l'effet réel du traitement. De nombreuses échelles de qualité ont été publiées (50), nous avons choisi la grille d'évaluation méthodologique de Jadad (9) car à notre connaissance il s'agit de la seule grille ayant été développée avec une méthodologie rigoureuse : 6 juges choisissaient chacun 49 items non redondants appliqués à 36 articles (6 étant jugés comme excellents, 6 étant jugés comme insuffisants par les grilles existantes et 23 articles étaient tirés au sort dans une banque de

données). Trente-huit items ont été retirés car non valides. Les scores pour les articles précédemment jugés comme excellents étaient significativement supérieurs aux scores obtenus par les articles tirés au sort ; ces 2 groupes d'articles obtenaient des scores significativement supérieurs au score obtenus par les articles précédemment jugés comme insuffisants. Cette grille a été développée et validée pour évaluer des ECR dans le traitement de la douleur, elle a largement été utilisée et validée dans d'autres type d'ECR (32).

L'intérêt de l'utilisation d'une grille d'évaluation méthodologique est débattue, et des études empiriques récentes (51,52) montrent qu'il est possible d'écarter la possibilité de biais avec suffisamment de certitude en utilisant seulement 3 critères (ce qui revient à utiliser le score de Jadad à 3 items).

- 1) Le caractère aléatoire de la répartition des patients dans les groupes traités et plus particulièrement son imprévisibilité qui garanti que les investigateurs ne pourront pas déterminer à l'avance dans quel groupe ira le patient qu'il souhaite inclure.
- 2) Le suivi en double insu si celui-ci est éthiquement possible.
- 3) L'absence ou un faible taux (<5%) de patients randomisés mais non inclus dans l'analyse (perdu de vue, analyse non réalisée suivant le principe de l'intention de traiter et excluant donc les arrêts prématurés de traitement, les écarts au protocole, etc.).

Lors de la réalisation d'une méta-analyse, il convient de vérifier que les critères de qualité méthodologique effectivement utilisés n'ont pas permis d'inclure des essais ne répondant pas à ces trois critères. De même, il est nécessaire de vérifier que le processus de sélection n'était pas trop strict et qu'il n'a pas entraîné l'exclusion d'essais dont la qualité était suffisante pour être inclus. L'utilisation d'une grille a l'avantage de transformer une donnée qualitative en une donnée quantitative permettant les comparaisons statistiques.

L'une des critiques que l'on peut faire à l'analyse des publications par grilles réside dans le fait que l'on est obligé de s'en tenir à ce qui est rapporté par les auteurs, qui ont pu être amenés à censurer leur manuscrit pour respecter les consignes rédactionnelles des éditeurs. Toutefois le lecteur et la communauté médicale ne disposent que des éléments publiés et l'auteur doit leur fournir les moyens de se faire une opinion objective.

Chalmers (53) a analysé le phénomène des données manquantes dans les publications en envoyant un questionnaire à 58 auteurs, 58% des données manquantes n'avaient pas été recueillies au sein de 41 essais dont les auteurs avaient renvoyé le questionnaire. On peut donc dire que l'on a, à peu près, 1 chance sur 3 de considérer à tort qu'un critère n'avait pas été rempli lorsqu'il n'était pas rapporté dans la publication (nous considérons ici que les données devaient être manquantes pour la majorité des essais dont les auteurs (n= 17) n'avaient pas renvoyé le questionnaire).

De nombreuses études ont mis en évidence les faiblesses méthodologiques des ECR publiés. Seulement 27% de 45 ECR publiés en 1985 avaient défini un critère d'évaluation (54), le calcul du nombre de patients nécessaire n'était calculé que parmi 43% des 47 ECR avec résultats négatifs publiés en 1990 (55). Certains essais rapportent aussi des résultats erronés, parmi 119 ECR rapportant des résultats en « intention de traitement » (lors de l'analyse, les patients restent dans leur groupe initial quel qu'est été le traitement effectué), 15 (13%) avaient exclu certains patients ou les avaient changé de groupe (56). Parmi 206 ECR de gynécologie-obstétrique décrits comme étant randomisés, 5% ne l'était pas (57). Des ECR de qualité méthodologique insuffisante ou rapportés de façon inadéquate peuvent être responsables de manquements dans la prise en charge de décisions sur le plan individuel mais aussi au niveau des politiques nationales de santé publique. Dans cette étude, 69 % (58/84) ECR avaient une qualité méthodologique satisfaisante, cette proportion a pu être surestimée en raison du choix des périodiques ou ont été sélectionnés les essais ; seulement 26 ECR (31%) avaient obtenus le score maximal vers lequel devrait tendre tous les ECR.

L'autre critique que l'on peut faire aux grilles d'analyses méthodologiques est que ces grilles s'appliquent à posteriori et n'ont de ce fait que peu d'utilité et d'impact dans l'amélioration des publications. Certaines études (58-60) ont mis en évidence une absence de l'amélioration de la qualité méthodologique des ECR au cours du temps.

Devant ce constat alarmant, un groupe de scientifiques et de rédacteurs de revues ont développés le « CONSORT » (consolidated standards of reporting trials) dans le but d'améliorer la qualité des publications des ECR. Ce groupe a établi une « check-list » et un organigramme (permettant facilement de comprendre le devenir de tous les patients de l'essai) que les auteurs peuvent utiliser lors de la rédaction de leur manuscrit. (annexe 5). CONSORT n'a pas comme but final l'évaluation des articles mais l'amélioration de la qualité des publications en apportant une aide, une trame aux auteurs ayant à rédiger la

publication d'un ECR. Le rapport CONSORT n'est pas figé et depuis sa première publication en 1996 (61), des mises à jour régulières ont été réalisées, la dernière version a été publiée en 2001 (62) ; un site Internet est également accessible (<http://www.consort-statement.org>) permettant de disposer de toutes les informations afin d'utiliser au mieux cette aide. L'utilisation du CONSORT est recommandée par un nombre croissant de revues (61-64) et par « The international committee of medical journals editors » (65) , plus connu sous le nom de « groupe de Vancouver ».

Une analyse de l'effet du CONSORT (66) sur la qualité méthodologique des ECR a été menée sur des ECR publiés en 1994 (avant publication du CONSORT) et en 1998 au sein de 4 revues : 3 ayant adopté et soutenu le CONSORT (BMJ, Lancet, JAMA) et le NEJM (n'ayant pas adopté le CONSORT). Le nombre d'éléments dont la présence était recommandée par le CONSORT étaient rapportés plus fréquemment en 1998 qu'en 1994 au sein des 4 revues, mais il existait une différence significative en défaveur du NEJM. La qualité méthodologique (score de Jadad) était améliorée pour les articles parus dans les 4 revues , mais il existait ici aussi une différence significative en défaveur du NEJM. Cette étude confirme l'efficacité de la méthode , mais aussi que les « cliniciens-chercheurs » prennent conscience de la nécessité d'améliorer la qualité méthodologique de leurs essais aussi bien dans la conception et la réalisation que dans la publication des résultats.

*"The RCT is a very beautiful technique, of wide applicability, but as with everything else there are snags. When humans have to make observations there is always the possibility of bias"*  
Cochrane AL (67)

Le score éthique global calculé pour les 84 ECR de l'étude était assez décevant :  $0.366 \pm 0.124$  . Ce score est inférieur au score obtenu dans l'étude de Berdeu et al (11) concernant un échantillon de 24 ECR ( $0.44 \pm 0.15$ ) mais supérieur au score obtenu par les ECR (68) sur les sujets âgés ( $0.334 \pm 0.118$ ) .

Dans cette étude l'obtention du consentement des patients n'est pas mentionné dans 2.4% (n = 2) des ECR étudiés, l'accord préalable d'un comité d'éthique n'est pas mentionné dans 21.4% (n =18) des cas et dans 5.9% (n =5 ) des cas, ni l'obtention du consentement des patients, ni l'accord préalable d'un comité d'éthique n'étaient mentionnés. Ruiz Canela et al (69) rapportent dans une étude portant sur 767 ECR publiés entre 1993 et 1995 (NEJM, Lancet, BMJ, JAMA) que 18.9% (n = 145) de ces ECR ne mentionnaient pas le recueil des consentements éclairés, 29.2% (n = 224) ne mentionnaient pas l'accord préalable d'un comité d'éthique et 12.9% (n = 99) ne mentionnaient ni l'un ni l'autre. Un questionnaire a été adressé aux auteurs dans le but de vérifier si les critères omis dans la publication originale avaient été ou non effectivement remplis. Sur 70% environ de questionnaires retournés, 22.4% des auteurs n'avaient pas demandé l'accord préalable d'un comité d'éthique et 20.6% des auteurs n'avaient pas recueilli le consentement éclairé. Les études menées par Ruiz- Canela (12,69) sont accablantes, montrant que plus de 60% (468/ 767) des auteurs ne respectent pas les critères éthiques légaux, ce pourcentage se réduit à 21.7% (58/267) lorsque l'on ne prends en compte que les auteurs ayant retournés le questionnaire. Cependant le fait de ne pas mentionner les critères éthiques réglementaire (même s'ils ont été recueillis) montre le manque d'intérêt de ces auteurs pour l'aspect éthique des ECR (70). Il est encore plus surprenant de constater que ces articles aient été acceptés pour publication dans des revues prestigieuses alors que les auteurs ne respectaient ni les personnes ayant participées au ECR ni la législation en vigueur.

Après ces constatations sur les manquements dans le respect des critères éthiques réglementaires ou légaux, il n'est pas surprenant d'observer que les autres critères éthiques soient aussi mis à mal. L'analyse critère par critère montrait que les ECR ne satisfaisaient régulièrement (plus de 70%) que les critères 1 , 2 , 4 et 7. Des analyses intermédiaires et des règles d'arrêt précoce n'étaient respectivement prévues que dans 31 et 19% des ECR de cette étude alors que les patients étaient soumis à des traitements lourds à toxicité importante. Nous avons mis en évidence que la qualité éthique ne s'améliorait pas avec le temps et n'était pas reliée à la revue dans laquelle l'ECR était publié. Par contre une

relation entre la qualité éthique et le nombre de patients inclus ont été mise en évidence, il existerait donc deux niveaux de qualité, en fonction de la taille de l'essai et donc un certain manque de rigueur pour les essais à petits effectifs. Contrairement à Berdeu et al (11), une relation a été mise en évidence dans cette étude entre la qualité éthique et la qualité méthodologique ( $p=0.006$  ,  $R^2=0.087$ ). Cette divergence entre les 2 études pourrait en partie s'expliquer par la différence du nombre d'ECR évalués dans ces deux études. Il semble toutefois que la qualité éthique d'une publication dépende essentiellement de l'intérêt qu'ont les concepteurs d'un essais vis a vis de la protection des personnes qu'ils soumettent à la recherche clinique.

L'amélioration de la qualité éthique des ECR est une nécessité, le respect et la protection des personnes devant être une préoccupation omniprésente, des actions peuvent être menées tout au long de la vie d'un essai c'est à dire de sa conception à sa publication. La formation des cliniciens-chercheurs se fait bien souvent sur « le tas », la sensibilisation et la réflexion sur les problèmes posés par l'expérimentation sur l'homme devraient faire partie de la formation initiale et continue de toute personne étant en relation avec la recherche (du chercheur au lecteur).

Le but du comité d'éthique ou du comité de protection des personnes est justement la protection des personnes, cette protection impose bien évidemment que le sujet ai reçu une information honnête, compréhensible et la plus complète possible avant de donner son consentement ; Que le rapport bénéfice risque soit en rapport du bénéfice et non du risque ; Que les garanties prévues pour les personnes qui se prêtent à la recherche aient bien été prévues... Mais si le rôle des comités est réduit à cela, une large proportion de patient acceptant de prendre le risque de se soumettre à une expérimentation ne sera pas protégée car ces personnes seront soumise à une expérimentation inutile. Une expérimentation peut être inutile pour plusieurs raisons :

- Les résultats de l'étude proposée sont déjà connus avec un degré de preuve suffisant
- L'étude ne pourra pas être menée à son terme
- La méthodologie n'est pas satisfaisante et les résultats obtenus ne seront pas exploitables
- Les résultats de l'étude ne seront pas publiés

Quel patient accepterait de prendre des risques en étant clairement informé que l'expérimentation qu'on lui propose est inutile ?

Le comité de protection des personnes doit avoir la compétence et les moyens de contrôler les essais du protocole à la publication , en s'assurant que l'équipe de chercheurs est qualifiée pour mener à bien le projet, que la qualité méthodologique est correcte, que le but de la recherche est valide, que la personne se prêtant à la recherche est protégée et enfin que l'étude sera utile à la communauté par sa diffusion. Ceci nécessite des moyens humains et financiers importants.

Les éditeurs de revues médicales ont un rôle important à jouer, en exigeant des garanties en ce qui concerne les précautions prises pour la protection des patients inclus dans les ECR. Certaines revues ont inclus dans leur « recommandations aux auteurs » (annexe 6), la nécessité d'un paragraphe consacré aux aspects éthiques. Cette pratique devrait se généralisée de même qu'une liste de contrôle (check-list) éthique permettant de vérifier que tous les critères attendus sont bien rapportés par l'auteur. Cette attitude a un double avantage, en premier lieu de ne publier que des études ayant respectées les malades , mais aussi d'être un instrument d'éducation et de sensibilisation des chercheurs ou des futurs chercheurs.

*« Add more  
life to your years  
rather than more  
years to your life »*  
Hopkins (71)

En cancérologie, l'efficacité thérapeutique à longterm était jugée en terme de survie ou de survie sans récurrence. La nécessité de s'attacher à la QDV des patients est la conséquence de plusieurs phénomènes. Pour certains types de cancer, des progrès spectaculaires en terme de survie ont été possible grâce à l'apparition de traitements de plus en plus efficace mais aussi de plus en plus toxiques (poly-chimiothérapie, association radiothérapie-chimiothérapie, greffe de moelle, immunothérapie...); Le thérapeute est souvent confronté à des effets secondaires parfois non-réversibles et parfois mortels. L'allongement de la durée de vie des patients et le suivi des patients guéris (40% selon Hill (72)) de leur cancer ont permis de mettre en évidence l'existence de séquelles physiques, psychologiques et sociales mais aussi de décrire les effets à long terme des traitements. Le thérapeute a donc aujourd'hui comme préoccupation de réduire la toxicité des traitements qu'il prescrit tout en conservant leurs efficacités

D'autres types de cancers n'ont pas bénéficiés des nouveaux traitements et le pronostic reste sombre. L'objectif du thérapeute n'est pas de proposer un traitement agressif à la recherche d'une guérison impossible mais d'évaluer de nouvelles stratégies thérapeutiques pouvant prolonger la survie sans grever la qualité de celle ci. Sans trop risquer de contradiction, on peut admettre que l'intérêt que chacun d'entre nous attache à la vie comporte deux éléments qui sont sa durée et sa qualité.

Il y a une quinzaine d'années, le « groupe de qualité de vie » de l'EORTC initiait un projet de développement d'une échelle utilisable pour toutes les localisations cancéreuse, l'actuel QLQC30. A la même époque Schraub et al (73) introduisaient en France l'évaluation de la QDV. La QDV est un concept global subjectif et multidimensionnel. Comme le bonheur la QDV est un concept équivoque (74-77) dont chacun possède une définition personnelle, intuitive, globale et difficilement communicable.

La première tentative de conceptualisation a eu comme point de départ la définition de la santé proposée par l’OMS en 1947 (78) : « *Un état de complet bien-être physique, mental et social, et non pas seulement l’absence de maladie ou d’infirmité* ».

Cette formule met l’accent sur la dimension subjective de l’état de santé et sur ces 3 principales composantes : physiques, psychiques et sociales. La qualité de vie est un concept intuitif général qui englobe l’état de santé, les caractères de l’environnement, le niveau de revenu économique, les conditions de vie affective, l’autonomie, les habitudes, le style de vie... Des définitions aussi larges dépassent de beaucoup ce qui est réalisable dans une étude clinique. Dans le domaine médical le concept de qualité de vie a été ciblé sur la qualité de vie liée à la santé (health related quality of life). Calman (79) a illustré sous forme de schéma toute la complexité de l’évaluation de la QDV et sa variation entre 2 individus face à un même problème mais aussi sa variation pour un même individu au cours du temps (annexe 7).

La satisfaction générale de la vie résultant de la différence entre les espoirs, les rêves, l’idéalisations de la vie et la réalité.

Malgré la complexité du concept de qualité de vie, seulement 12% des auteurs donnaient leur définition – conception de la qualité de vie qu’ils souhaitaient mesurer et seulement 54.5% des auteurs indiquaient quels domaines de la qualité de vie ils avaient mesuré.

L’appréciation de la QDV individuelle peut être réalisée par un entretien psychologique mais il s’agit d’une approche difficile à intégrer dans une étude en raison de son coût et des problèmes liés à la reproductibilité.

Des échelles de mesure ont donc été développées, un bon outil de mesure doit être :

- pertinent, c’est à dire mesurer de façon valide ce qu’il est supposé mesurer
- performant, c’est à dire posséder de bonnes qualités métrologiques
- applicable facilement en pratique (temps de réponse court)

On distingue cinq types de validité :

- La validité perçue : le degré d’acceptabilité doit être bon (la compréhension des questions, le caractère intrusif de certaines questions, le temps de remplissage). La validité perçue est évaluée par les données manquantes et les commentaires du patient à la fin du questionnaire.
- La validité du contenu (spécificité) est la capacité de l’échelle à mesurer ce qu’elle est censée mesurer. La validité de contenu est évaluée en faisant appel à des experts.

- La validité de structure est la capacité de l'échelle à mesurer les différents types de fonctionnements retenus (physiques, psychique, social). Elle est proche de la validité de contenu.
- La validité de concordance est la corrélation entre une dimension donnée de l'échelle avec une échelle extérieure validée mesurant cette même dimension.
- La validité de prédiction est la capacité à prédire des événements ultérieurs. Un score peu élevé est-il prédictif d'une survie courte ?

Les qualités métrologiques sont au nombre de 3 :

- La sensibilité (responsiveness) : instrument fournissant des résultats différents d'un sujet à l'autre ou pour un même sujet au cours du temps.
- La spécificité : l'instrument ne mesure que ce qu'il est censé mesurer
- La fiabilité : 2 mesures successives fournissent des résultats comparables.

La plupart des outils de mesure disponibles sont des questionnaires composés de réponses fermées : réponses dichotomiques (oui/non), réponses qualitatives ordonnées (score de 1 à 5), réponse visuelle analogique.

### **Comment choisir l'instrument de mesure ?**

Il existe de très nombreux instruments de mesure, mais le questionnaire idéal n'existe pas. Un des éléments à prendre en compte est d'établir la liste des pays participants à l'étude afin de choisir un instrument traduit et validé pour le plus grand nombre de pays participants.

L'instrument utilisé doit être applicable dans la pratique : longueur du questionnaire, temps de remplissage tenant compte de l'état de fatigue des patients, compréhension des questions sur le plan culturel ou en présence d'éventuels troubles cognitifs. Toutefois, le principal objectif à considérer est l'objectif de l'étude afin de déterminer quels aspects de la QDV sont pertinents pour y répondre. En cancérologie, il peut sembler qu'un questionnaire spécifique soit le mieux adapté surtout lorsque l'on cherche (et c'est le cas des ECR) à évaluer l'importance des changements survenus chez un individu ou un groupe d'individus. Avant de faire ce choix, il ne faut pas oublier que des paramètres de santé globale peuvent interférer peut-être de façon différente selon les traitements ; dans ce cas seul un questionnaire générique les prendra en compte.

Parmi les questionnaires disponibles en cancérologie, le QLQC30 et le FACT sont constitués d'un questionnaire central, général pour tout cancer, et des modules spécifiques de différents organes ou symptômes. L'intérêt de cette structure modulaire est manifeste, elle permet d'éviter de multiplier les questionnaires pour chaque type de cancer. Les modules spécifiques explorent à la fois les retentissements de la maladie que les effets secondaires des traitements. Si des points particuliers de l'étude n'étaient pas mesurés par l'instrument chois, il est possible d'ajouter des questions (80) qui seront analysées à part en sachant que leur fiabilité et leur validité n'ont pas été validées. L'organisation en modules de ces questionnaires prend peut-être part à la prépondérance de l'utilisation de ces deux questionnaires. Cependant 13 instruments différents ont été utilisés au sein des 72 ECR comportant une analyse de la QDV. Cette multiplicité des questionnaires utilisés n'est pas sans poser un problème, en effet Kemmler et al (81) dans une étude portant sur 244 patients ont mis en évidence que malgré des redondances, une seule des 5 dimensions du FACT, la dimension physique, était bien explorée par le QLQC30. Seulement trois (physique, état de santé global, symptômes généraux) des huit dimensions du QLQC30 étaient bien explorées par le FACT. Cette étude montre qu'il n'est pas possible de remplacer l'un par l'autre dans une même étude et que la comparaison des résultats obtenus par les 2 instruments est impossible.

La non comparabilité des résultats de ces 2 questionnaires (et des autres) posera certainement un problème, lorsque la qualité de vie sera prise en compte des méta-analyses ou des SOR, sans oublier que 12 questionnaires de qualité de vie différents ont été utilisés ainsi que dans 4 cas des questionnaires « développés » uniquement pour l'étude.

Une des solutions serait de développer un nouvel instrument de mesure combinant des éléments du FACT et du QLQC30 mais si cet instrument était validé l'un des risques serait de voir l'instrument s'imposer pour toutes les études et comme dans beaucoup de domaines la diversité est une richesse. Il semble être préférable de poursuivre la recherche en QDV sur des voies différentes plutôt que de chercher l'instrument parfait.

### **Qui doit effectuer la mesure ?**

Il est admis maintenant par tous que c'est le sujet qui est le plus à même de parler de ce qu'il vit (82) et d'exprimer avec toute sa subjectivité le décalage entre ce qu'il souhaite et ce qu'il perçoit (77). Les instruments de mesure dont nous disposons sont en fait des auto-questionnaires, dans cette étude un seul questionnaire complété par le corps médical (Spitzer) a été utilisé et il était couplé à un auto-questionnaire.

Des études expérimentales (25, 83, 84) confirment que le corps médical est un mauvais juge de la QDV des patient, le taux de corrélation entre la QDV estimée par les médecins ou les patients allait de 0.27 à 0.63. Dans ces trois études le score de qualité de vie était plus haut lorsque l'évaluation était faite par les malades.

Kahn et al (85) ont mis en évidence dans une étude appariée que les médecins évaluaient de façon erronée la QDV des patients jeunes. Les médecins estimaient que les patients âgés présentaient une altération plus importante que les jeunes dans les domaines suivants : physiques, financiers, sociaux et médicaux. Les auto-questionnaires ont révélé que les patients jeunes avaient une altération de la QDV plus importante que les patients âgés.

L'importance de l'évaluation par auto-questionnaire est souligné par des résultats non attendus et surprenants comme l'étude de Sugurbaker et al (86).

Cette étude étudiait la QDV des patients porteurs de sarcomes des extrémités. L'hypothèse de l'auteur était qu'une chirurgie conservatrice associée à une radiothérapie altérerait moins la QDV qu'une chirurgie mutilante. Les résultats ont montré que les patients amputés rapportaient moins de difficultés dans les activités quotidiennes, pour les soins corporels et les déplacements ; moins de troubles de l'émotion et moins de problèmes dans leur sexualité que les patients ayant « bénéficié » du traitement conservateur.

### **Quand doit-être administré le questionnaire ?**

La qualité de vie est un phénomène évolutif et continu, la mesurer à des moments définis risque d'entraîner des biais dans son évaluation et son interprétation. En pratique il est impossible d'effectuer une mesure en continu. La multiplication des mesures risque d'entraîner un refus des patients de continuer à compléter les questionnaires.

La première mesure, mesure de référence, doit être réalisée avant la randomisation ou au plus tard avant que soit donné au patient l'information sur les modalités du traitement qu'il va recevoir.(87). En effet les réponses données après randomisation pourraient être influencées par la connaissance par le patient du traitement qu'il va recevoir . Si cette

mesure de référence n'est pas faite, aucune comparaison ne pourra être effectuée lors de l'analyse des résultats.

Par la suite les investigateurs devront utiliser le moins de mesure possible afin de ne pas surcharger le patient, la connaissance de l'histoire naturelle de la maladie et du traitement aideront les investigateurs à choisir quelles mesures sont pertinentes et nécessaires. Les mesures doivent être planifiées afin de permettre une comparaison honnête entre deux traitements.

#### Les mesures sont planifiées en fonction des buts de l'étude.

-si l'on veut mesurer l'effet de la toxicité du traitement sur la qualité de vie, la mesure devra avoir lieu au moment où la toxicité sera maximale.

-si l'on veut mesurer l'effet du traitement sur les symptômes de la maladie altérant la QDV, la mesure devra avoir lieu lorsque la toxicité du traitement sera estompée ; juste avant le début d'une nouvelle cure dans le cas d'une chimiothérapie.

Les mesures doivent se poursuivre jusqu'à la fin de l'étude pour permettre une comparaison correcte entre les deux groupes sans tenir compte du fait que dans un groupe le traitement a pu être arrêté (par exemple : 3 cures de chimiothérapie versus 6 cures). Être sans traitement ne signifie pas être hors de l'étude.

Le problème du moment où doit être administré le questionnaire est un problème réel comme le témoigne les résultats de cette étude où 15 fois le moment de la mesure n'était pas précisé et pour les 18 autres ECR le remplissage se faisait avant, pendant, après la cure ou quand le désirait le patient. Nous avons vu que la toxicité devait être différenciée de la QDV, ceci souligne l'intérêt des échelles de toxicité (utilisée dans 98% des ECR de cette étude) qui ne peuvent ni remplacer ni être remplacées par des questionnaires de QDV, certaines études (ont mis en évidence qu'une augmentation de la toxicité du traitement pouvait se traduire par une amélioration de la QDV si la maladie régressait (26, 82, 89).

#### **Pour quel type d'essais doit-on mesurer la qualité de vie ?**

Osoba et al (4), considèrent que l'étude de la QDV devrait être incluse pour tous les essais de phase III en cancérologie, et il propose même (90) que la mesure de la QDV devienne le « gold standard » auquel les autres variables devraient être comparées.

Inclure une mesure de la QDV au sein des essais, augmente de façon considérable le travail de recueil des données, la gestion et l'analyse de celles-ci. Ce que l'on attend des résultats est-il suffisant pour motiver les différents acteurs (dont le patient) tout au long de l'étude et de supporter l'augmentation de ses coûts en temps et en argent. La remise d'un

questionnaire de QDV n'est pas une démarche anodine, de nombreux patients remplissent des questionnaires en espérant avoir un retour personnel sur la souffrance psychosociale qu'ils expriment dans ce questionnaire. Or ce retour n'est généralement pas donné. Les questions interrogent le patient sur sa vie personnelle et certaines d'entre elles peuvent être perçues comme indiscrètes ou génératrices d'angoisse faisant prendre conscience au patient d'un état qu'il refuse d'accepter.

L'évaluation de la QDV ne peut constituer une mesure de routine dans tous les essais ; les résultats qu'on en attend doivent apporter une aide réelle à la décision dans la prise en charge du malade, qu'elle soit thérapeutique, psychologique ou sociale.

Il ne s'agit pas de répondre à un phénomène de mode ni de rechercher un label qualité en utilisant un questionnaire de QDV dans un essai thérapeutique. Dans tous les cas cette inclusion doit profiter au malade et, comme le dit Pocock (91) avant d'inclure la mesure de qualité de vie dans cette étude, il faut d'abord se demander si, tout simplement, ce n'est pas « pour mettre la cerise sur le gâteau » (« it is not icing on the cake »).

Nous essayerons ici de déterminer quels essais thérapeutiques peuvent être optimisés par l'étude de la QDV ; Nous ne développerons pas ici l'intérêt de l'évaluation de la QDV dans l'information faite aux patients concernant les effets secondaires des médicaments (92) ni de l'intérêt de la mesure de la QDV avant traitement comme facteur pronostic des cancers. (93)

Nous nous limiterons donc au cadre des essais thérapeutiques en nous inspirant des réflexions de différents auteurs (94-98) et l'EORTC (99)

\* Essais où la QDV est le « primary end-point » : Par exemple essais comparant un traitement par chimiothérapie à des soins de confort, d'accompagnement (communément qualifié de soins palliatifs, mais la chimiothérapie a-t-elle une prétention curative ?). Comparaison de deux traitements palliatifs.

\* Essais d'équivalence ou de tentative de desescalade des traitements, on attend une diminution de la toxicité avec conservation de l'effet sur la survie...

Comparaison de deux traitements dont la différence d'effet en terme de survie, est faible mais dont la différence de toxicité est élevée.

\* Essais comparant deux traitements dont l'effet en terme de survie sera important mais au prix d'une toxicité élevée pour l'un des traitements (par exemple : chimiothérapie standard vs dose élevée de chimiothérapie suivie de greffe de moelle osseuse).

\* Essais où l'effet des traitements va être faible sur la survie à long terme et où la médiane de survie sera courte ; essais portant sur des stades avancés de cancer (cadre de l'étude).

\* Essais de « coûts-efficacité » : la limitation des ressources pouvant être attribuées aux actes de santé impose une rationalisation des dépenses tout en répondant à la demande de qualité des soins et de la vie des individus.

### **Le problème posé par les données manquantes**

Les données manquantes représentent le principal biais des études portant sur le QDV.

Dans cette étude le taux d'observance vis à vis du remplissage des instruments de mesure de la QDV allait de 26% à 100%.

Le défaut d'observance peut intervenir à 4 niveaux :

- Au niveau de l'essai multicentrique, certains centres peuvent décider de ne pas participer à l'évaluation de la QDV (100).
- Au niveau d'un des centres : certains patients acceptent de participer à l'évaluation de la QDV d'autres non.
- Au niveau du patient : un patient peut compléter un questionnaire à un moment t et ne pas le compléter au moment t'. (problème de fatigue, de la baisse de l'état de santé rendant le remplissage du questionnaire impossible).
- Au niveau du questionnaire : certaines questions auront une réponse alors que d'autres resteront sans réponse.

De plus à tous les niveaux peuvent intervenir les erreurs, oublis.....

Gérer les données manquantes posent des problèmes au stade de l'analyse et de l'interprétation (101-102).

Un taux de réponse de 60% pour les études avec envoi par voie postale est considéré comme un excellent taux de réponse, ce pourcentage est considéré comme mauvais dans les études de QDV. Les méthodes utilisées pour estimer les données manquantes ne sont actuellement pas satisfaisantes (103)

En effet une étude menée par Brorsson et al (104) en suède, en recontactant par téléphone les patients (40%) n'ayant pas complété les questionnaires, a montré que les non-répondeurs avaient une meilleure QDV que les répondeurs. Des résultats inverses ont été rapportés où les non-répondeurs avaient une baisse de la QDV le défaut d'observance était due à une toxicité, une morbidité ou une mortalité plus importante.(105)

Une étude (106) a montré que lorsque le pourcentage de donnée manquante était supérieur à 25% des biais dans l'interprétation des résultats existaient quel que soit les techniques mathématiques ou statistiques utilisées.

Le protocole d'un essai doit avoir prévu la méthode pour gérer les données manquantes ; dans cette étude seulement 3% des ECR rapportaient cette information.

Un autre problème posé par l'estimation et la comparaison de la QDV rejoignant le problème des données manquantes est celui du calcul du nombre nécessaire de patient à inclure dans l'essais. En effet si l'on évalue la QDV par 1 unique score global, le calcul se fera selon la procédure commune à tous les essais. Mais dans la plupart des essais, la comparaison se fait sur plusieurs dimensions ou domaines (social, physique...), il faudrait donc calculer le nombre de patients nécessaire pour chacune des dimensions évaluées. En pratique le nombre de patients nécessaires est calculé sur un autre critère que celui de la QDV.

### **Problèmes liés à l'interprétation des résultats**

Le résultat d'un questionnaire de QDV peut-être exprimé par un score global ou un score pour les différents domaines étudiés.

Lorsqu'un résultat est exprimé par un score global, la comparaison est réalisée par un test unique. Par contre lorsque les résultats sont donnés par un score par dimension étudiée il y a autant de test d'hypothèses que de dimensions étudiées et le risque de conclure à tort (risque de 1<sup>er</sup> espèce) augmentent avec le nombre de tests. L'usage d'un score unique augmente la probabilité de détecter une différence alors que les différences pour chaque domaine étudié sont minimales (107).

L'usage d'un score unique peut permettre de détecter une différence mais ne permet pas de savoir pourquoi cette différence existe.

D'autre part, l'usage d'un score par domaine étudié, peut rendre l'interprétation difficile si les différentes comparaisons ne vont pas dans le même sens.

Les résultats de cette étude montre la difficulté de l'interprétation et de la discussion des variations de la QDV. Dans 27.3% des ECR les auteurs ne discutaient pas sur les variations ou l'absence de variation de la QDV mise en évidence dans leur étude.

L'utilisation des instruments de mesure de la QDV semble être considérée par les auteurs(même si l'usage se développe) comme un critère secondaire comme en témoigne le

pourcentage élevé d'article ou la mention de la mesure de la QDV n'était pas faite au niveau de l'abstract.

Une attitude sage à peut-être été prise par certains auteurs (n=5\*), la publication des résultats de la variation de la QDV dans un article séparé. Notre préférence va à la publication succincte des variations de QDV dans l'article initial et une description complète dans un 2<sup>ème</sup> article (n=3).

Cette attitude n'est pas innocente et a bien sûr pour les auteurs, l'avantage de publier deux articles à propos du même essai. Pour les lecteurs l'avantage est d'avoir un rapport complet sur les variations de la QDV ce qui ne serait pas possible dans un seul article en raison de l'hypertrophie de l'article dont serait responsable le rapport détaillé de l'essai en terme de réponse tumorale ou de survie et en terme de QDV.

(\*Les 5 articles attendus sur la QDV n'étaient pas encore publiés lorsque la présente étude a été réalisée).

### **problèmes éthiques liés à l'utilisation des questionnaires**

l'acceptabilité des questionnaires de qualité de vie est souvent ignorée tant au niveau de leur conception ou de leur traduction que de leur utilisation(6-8).

Problématiques liées à *l'absence de retombées pour le patient*. En effet de nombreux patients remplissent des questionnaires de qualité de vie lors d'essais thérapeutiques en espérant avoir un retour personnel sur la souffrance psychosociale qu'ils expriment dans ces questionnaires. Or ce retour n'est généralement pas donné, l'utilisation des questionnaires de qualité de vie s'intégrant uniquement dans le cadre d'un essai thérapeutique. L'analyse des questionnaires est faite de façon tardive avec une évaluation globale sans bénéfice individuel.

Problématique *spirituelle*. Certaines questions formelles risquent de ramener à la surface des sujets tabous enfouis dans l'inconscient et risquent de ce fait de générer une anxiété en faisant prendre conscience au patient d'un état qu'il refusait d'accepter. (« Je suis préoccupé (e) par l'idée de mourir » : je vais mourir, j'ai donc une maladie mortelle....., « j'ai peur que mon état s'aggrave ».....)

Problématique liée à *l'agression, l'injure de certaines questions* non adaptées à l'état du patient. (« je me sens capable de travailler » : mais docteur vous voyez bien que je ne suis pas capable de travailler, je ne peux même plus marcher...) Une interrogation sur la longueur d'une marche qu'un malade peut faire est mal ressentie par un patient métastatique alité depuis de nombreuses semaines.

Problématique liée aux *questions gênantes, intimes* (« Avez vous eu une vie sexuelle au cours de l'année passée ? Si oui je suis satisfait (e) de ma vie sexuelle »)

Problématique liée aux *conséquences éventuelles de certaines questions*. : « j'ai confiance en mon (mes) médecin (s) », « mon (mes) médecin(s) est (sont) disponible(s) pour répondre à mes questions »

Ces questions feront l'objet de recherches tant est grand le domaine et l'application de la QDV.

## **Conclusion**

Les essais randomisés de phase III en cancérologie qui ont été analysés avaient une qualité méthodologique satisfaisante dans 69% des cas. La qualité éthique des ces essais était décevante avec environ 30% des auteurs qui ne respectaient pas les critères réglementaires et légaux (consentement éclairé et accord préalable d'un comité d'éthique).

Des actions de formation et de sensibilisation des chercheurs sont nécessaires, ainsi que le renforcement de la « surveillance » des essais par les comités de protections des personnes . les éditeurs de revues médicales sont également impliqués dans cette démarche de qualité et leur niveau de vigilance doit être revu à la hausse. Afin d'améliorer la qualité tant méthodologique qu'éthique, l'exemple CONSORT mérite d'être suivi et la même proposition pourrait-être faite sur le plan de l'éthique, ceci permet de passer d'une analyse à posteriori, au stade de l'action a-priori.

L'évaluation de la QDV dans les essais de cancérologie est en augmentation constante depuis le début des années 80. Cependant, le mode d'emploi doit-être défini de façon précise car les disparités observées dans l'utilisation des instruments de mesure rendent toute tentative d'analyse dans tel ou tel cancer illusoire. L' "evidence based quality of life » est encore loin.

Toutefois ,malgré ces critiques, l'évaluation de la QDV nous apporte des connaissances importantes. L'expérimentation sur l'homme change peu à peu de visage (visage humain), et l'évaluation de la QDV étant de plus en plus souvent prise en compte.

Une évolution qui a permis la mesure de la QDV est la modification du paradigme médical habituellement centré sur la maladie vers un paradigme médical centré sur la personne malade.

La QDV devient un critère important en dehors de la recherche clinique par son utilisation « quotidienne » permettant une meilleure prise en charge des patients. L'utilisation courante avec analyse des résultats par un système informatique pourrait être utile dans la décision thérapeutique au lit du malade.

## Références :

1. Karnofsky DA, Burchenal JH. The clinical evaluation of chemotherapeutic agents in cancer. In evaluation of chemotherapeutic agents, McLeod CM eds. Columbia University Press, New-York 1949; 191-205.
2. American Society of Clinical Oncology. Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. *J Clin Oncol* 1996; 14 : 671-9.
3. Whoqol Group. Study protocol for the World Health Organisation project to develop a quality of life assessment instrument (Whoqol). *Qual Lif Res* 1993;2: 153-9.
4. Osoba D. The quality of life committee of the clinical trials group of the national cancer institute of Canada: organisation and functions. *Quality of life research* 1992; 1: 211-8.
5. Gotay CC, Korn EL, McCabe MS, Moore TD, Cheson BD. Quality-of-life assessment in cancer treatment protocols: research issues in protocol development. *Journal of the National Cancer Institute* 1992; 84: 575-9.
6. Aaronson NK, quality of life and clinical trials. *Lancet* 1995; 346: 1-2.
7. Hewitt P, Dickersin K, Chalmers TC. More on MEDLINE searches. *Control Clin Trials* 1988 Mar;9(1):85-90
8. Adams CE, Lefebvre C, Chalmers I. Difficulty with MEDLINE searches for randomised controlled trials *Lancet* 1992 Oct 10;340(8824):915-6
9. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996 Feb;17(1):1-12
10. National Cancer Institute : quality of life in clinical cancer trials. Proceeding of a work-shop held at the NIH Bethesda (march 1-2, 1995) *JNCI Monographs* 1996, n° 20.
11. Berdeu D, Alric R, Fourcade J, Herve C. Qualité éthique et qualité méthodologique des essais contrôlés randomisés, à partir des études publiées de 1989 à 1992, dans quatre revues internationales. *Les cahiers d'Éthique Médicale, collection l'Éthique en mouvement* (c. Herve , Ed), 1/98, p 117-37, Editions l'Harmattan, Paris.
12. Ruiz-Canela M, de Irala-Estevez J, Martinez-Gonzalez MA, Gomez-Gracia E, Fernandez-Crehuet J. *J Med Ethics* 2001 Jun;27(3):172-6 Methodological quality and reporting of ethical requirements in clinical trials.

13. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994 Aug 24-31;272(8):619-26
14. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27 (suppl): S217- S232
15. Gertz MA, Lacy MQ, Lust JA, Greipp PR, Witzig TE, Kyle RA. Prospective randomized trial of melphalan and prednisone versus vincristine, carmustine, melphalan, cyclophosphamide, and prednisone in the treatment of primary systemic amyloidosis. *J Clin Oncol* 1999; 17: 262-7.
16. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993 Mar 3;85(5):365-76
17. Cella DF, Tulskey DS, Gray G, Sarafian B, Linn E, Bonomi A, Silberman M, Yellen SB, Winicour P, Brannon J, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993 Mar;11(3):570-9
18. De Haes JCJM, Olschewski M, Fayers P, et al. Measuring the quality of life of cancer patients with the Rotterdam Symptom Checklist: a manual. Groningen, Northern Centre for Healthcare Research, 1996.
19. Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients: the Functional Living Index-Cancer: development and validation. *J Clin Oncol* 1984 May;2(5):472-83
20. Mercier M, Bonnetterre J, Schraub S, Lecomte S, el Hasnaoui A. The development of a French version of a questionnaire on the quality of life in cancerology (Functional Living Index-Cancer: FLIC). *Bull Cancer* 1998 Feb;85(2):180-6
21. Hollen PJ, Gralla RJ, Kris MG, Potanovich LM. Quality of life assessment in individuals with lung cancer: testing the Lung Cancer Symptom Scale (LCSS). *Eur J Cancer* 1993;29A Suppl 1:S51-8
22. Ware JE, Kosinski M, Keller SD. SF 36 physical and mental summary scales: a user's manual. Boston, Ma, the health Institute 1994.
23. Sloan JA, Loprinzi CL, Kuross SA, et al. A randomized comparison of four tools measuring overall quality of life in patients with advanced cancer. *J Clin Oncol* 1998; 16: 3662-73.

24. Coates A, Fischer Dillenbeck C, Mc Neil DR, et al. On the receiving end-II. Linear Analogue Self-assessment (LASA) in evaluation of aspects of quality of life of cancer patients receiving therapy. *Eur J Cancer Clin Oncol* 1983; 19: 1633-7.
25. Spitzer WO, Dobson AJ, hall J, et al. Measuring the quality of life in cancer patients: a concise QL-index for use by physicians. *J Chron Dis* 1981; 34: 585-97.
26. Priestman TJ, Baum M. Evaluation of quality of life in patients receiving treatment for advanced breast cancer. *Lancet* 1976; 1: 899-901.
27. Kurihara M, Nakamura H, Matsukawa M, et al. Assessment of QOL in cancer drug therapy using 22-item questionnaire. *Gan To Kagaku Ryoho* 1994; 21: 379-87. (in japanese)
28. Hurny C, Bernhard J, Bacchi M, van Wegberg B, Tomamichel M, Spek U, Coates A, Castiglione M, Goldhirsch A, Senn HJ, et al. The Perceived Adjustment to Chronic Illness Scale (PACIS): a global indicator of coping for operable breast cancer patients in clinical trials. Swiss Group for Clinical Cancer Research (SAKK) and the International Breast Cancer Study Group (IBCSG). *Support Care Cancer* 1993 Jul;1(4):200-8
29. Mackowiak JI, Scott-Lenox J, Lennox R, et al. Patients benefit questionnaire (PBQ) for xerostomia: development and validation report. *Value Health* 1999; 2: 197 (abstract).
30. Dickersin K. The existence of publication bias and risk factors for its occurrence. *Jama* 1990;263(10):1385-9.
31. Cucherat M. Méta-analyse des essais thérapeutiques: Masson; 1997
32. Clark HD, Wells GA, Huet C, McAlister FA, Salmi LR, Fergusson D, Laupacis A. Assessing the quality of randomized trials: reliability of the Jadad score. *Controlled Clinical Trials* 1999;20: 448-52.
33. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. Does masking author identity improve peer review quality? A randomized controlled trial. PEER Investigators. *JAMA* 1998;280(3):240-2.
34. Fisher M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review. *Jama* 1994;272(2):143-6.
35. Mariani L, Marubini E. Content and quality of currently published phase II cancer trials. *J Clin Oncol* 2000 ; 18 : 429-36
36. Petticrew M. Diagoras of Melos (500 BC): an early analyst of publication bias. *Lancet* 1998;352(9139):1558

37. Cowley AJ, Skene A, Stainer K, Hampton JR. The effect of lorcinide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias. *Int J Cardiol* 1993;40(2):161-6
38. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324(12):781-8.
39. Moore TJ. *Deadly medicine*: Simon & Schuster; 1995
40. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337(8746):867-72
41. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *Jama* 1992;267(3):374-8.
42. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *Br Med J* 1997;315(7109):640-5.
43. Djulbegovic B, Lacey M, Cantor A, Fields KK, Bennett CL, Adams JR, et al. The uncertainty principle and industry-sponsored research. *Lancet* 2000;356(9230):635-8.
44. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350(9074):326-9
45. Huston P, Moher D. Redundancy, disaggregation, and the integrity of medical. *Lancet* 1996;347(9007):1024-6
46. Rennie D. Fair conduct and fair reporting of clinical trials. *Jama* 1999;282(18):1766-8
47. Do oncologists have an increasing interest in the quality of life of their patients ? a literature review of the last 15 years. *Eur J Canc* 1997; 33:29-32.
48. Sanders C, Egger M, donovan J, Tallon D, Frankel S. Reporting on quality of life in randomized controlled trials: bibliographic study. *Br Med J* 1998; 317: 1191-4.
49. Schraub S, Mercier M. Qualité de vie en cancérologie. *Bull Cancer* 2000 ; 87 : 117-20.
50. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Controlled Clinical Trials* 1995;16:62-73.

51. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
52. Moher D, Ba'Pham, Jones A, Cook DJ, Jadad AR, Moher M, Tugwell P, Klassen T. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses ? *The Lancet* 1998;352:609-13
53. Chalmers TC, Smith H, Blackburn B et al. A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials* 1981; 2: 31-79.
54. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med*. 1987;317:426-32.
55. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA*. 1994;272:122-4.
56. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*. 1999;319:670-4.
57. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA*. 1994;272:125-8.
58. Nicolucci A, Grilli R, Alexanian AA, Apolone G, Torri V, Liberati A. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. *JAMA*. 1989;262:2101-7.
59. Ah-See KW, Molony NC. A qualitative assessment of randomized controlled trials in otolaryngology. *J Laryngol Otol*. 1998;112:460-3.
60. Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*. 1998;317:1181-4.
61. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ*. 1996;313:570-1.
62. Moher D, Schulz KF, Altman DG, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports parallel group randomized trials. *Lancet*. 2001;357 : 1191-4.
63. Anonymous[A standard method for reporting randomized medical scientific research; the 'Consolidation of the standards of reporting trials' (CONSORT)] Verslaglegging van gerandomiseerd medisch-wetenschappelijk onderzoek volgens een standaardmethode; de 'Consolidation of the standards of reporting trials' (CONSORT). *Ned Tijdschr Geneeskd*. 1998;142:1089-91.

64. Huston P, Hoey J. CMAJ endorses the CONSORT statement. CONSolidation of Standards for Reporting Trials. CMAJ. 1996;155:1277-82.
65. Davidoff F. News from the International Committee of Medical Journal Editors. Ann Intern Med. 2000;133:229-31.
66. Moher D, Jones A, Lepage L; CONSORT Group. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 2001; 285(15):1992-5
67. Cochrane AL. Effectiveness and efficiency; random reflections on health services. Abingdon, Berkshire, UK: The Nuffield Provincial Hospitals Trust; 1972:2.
68. Berdeu D, Herve C, Fourcade J. Clinical trials in the elderly: ethical and methodologic considerations. Rev Med Interne 2000;21(7):614-22
69. Ruiz-Canela M, Martinez-Gonzalez MA, Gomez-Gracia E, Fernandez-Crehuet J. Informed consent and approval by institutional review boards in published reports on clinical trials. N Engl J Med 1999 ; 340(14):1114-5
70. Rennie D, Yank V. Disclosure to the reader of institutionnal review board approval and informed consent. JAMA 1997; 277: 922-3.
71. Hopkins A. Measuring the quality of medical care. Londres: Royal College of Physicians of London.1990.
72. Hill C. Quels taux de guérison pour le cancer. Bull Cancer 1998 ; 85 : 745-6.
73. Schraub S, Mercier M, Turkeltaub E, Fournier J. Mesure de la qualité de vie. Bull Cancer 1987 ; 74 : 297-305.
74. Cella DF. Quality of life outcomes : measurement and validation. Oncology 1996; nov. (suppl) 233-46.
75. Cohen SR, Mount BM, Mac Donald N. Defining quality of life. Eur J Cancer 1996; 32A: 753-4.
76. Fagot-Largeault A. Reflexion sur la notion de qualité de la vie. Archives de philosophie du droit1991 ; 36, droit et science : 135-53.
77. Leplège A. Mesurer la qualité de vie du point de vue des patients. L'enseignement philosophique 1996 ; 4 : 23-37.
78. OMS : The contribution of the world health organisation. I: Word Health Organisation. Who Chron 1947; 1 :29.
79. Calman KC. Quality of life in cancer patients – an hypothesis. Journal of Medical Ethics 1984; 10: 124-7.

80. Tannock IF, Osoba D, Stockler MR, et al. Chemotherapy with mitoxantrone plus prednisone or prednisone alone for symptomatic hormone-resistant prostate cancer: a Canadian randomized trial with palliative end points. *J Clin Oncol* 1996 Jun;14(6):1756-64.
81. Kemmler G, Holzner B, Kopp M, et al. Comparison of two quality-of-life instruments for cancer patients: the functional assessment of cancer therapy-general and the european organization for research and treatment of cancer quality of life questionnaire-C30. *J Clin Oncol* 1999; 17: 2932-40.
82. Dazord A, Guerin P, Brochier C, Cluse M, Terra JL, Seulin C. Un modèle de qualité de vie subjective adaptée aux essais thérapeutiques : intérêt chez les patients dépressifs. *Santé Mentale au Québec XVIII*, 2 : 49-74.
83. Presant CA. Quality of life in cancer patients. Who measures what? *Am J Clin Oncol* 1984; 7: 571-3.
84. Slevin ML, Plant H, Lynch D, et al. Who should measure quality of life, the doctor or the patient? *Br J Cancer* 1998; 57: 109-12.
85. Kahn SB, Houts PS, Harding SP. Quality of life and patients with cancer: a comparative study of patient versus physicians perceptions and its implications for cancer education. *J Cancer Educ* 1992; 7: 241-9.
86. Sugarbaker PH, Barofsky I, Rosenberg SA, et al. Quality of life assessment of patients in extremity sarcoma clinical trials. *Surgery* 1981; 9: 17-23.
87. Osoba D. Rationale for timing of health-related quality of life assessments in oncological palliative therapy. *Cancer Treat Rev* 1996; 22 (suppl A), 69-73.
88. Baum M, Priestman T, West RR, et al. A comparison of subjective responses in trial comparing of subjective responses in a trial comparing endocrine with cytotoxic treatment in advanced carcinoma of the breast, in Mouridsen HT, Palshoff ST (eds): *Breast cancer. Experimental and clinical aspects*. Oxford, England, Pergamon Press, 1980, pp 223-6.
89. Coates A, GebSKI V, Bishop JF, et al. Improving the quality of life during chemotherapy for advanced breast cancer. A comparison of intermittent and continuous treatment strategies. *N Engl J Med* 1987; 317: 1490-5.
90. Osoba D. What has been learned from measuring health-related quality of life in clinical oncology. *Eur J Cancer* 1999, 35 : 1565-70.

91. Pocock SJ. A perspective on the role of quality of life assessment in clinical trials. *Control Clin Trials* 1991; 2: 257S-65S.
92. Reynolds P, Sanson-Fisher R, Poole d, Harker J, Byrne M. Cancer and communication : information-giving in an oncology clinic. *Br Med J* 1981; 282: 1449-51.
93. Ganz R, Lee J, Siau J. Quality of life assessment: an independent pronostic variable for survival in lung cancer. *Cancer* 1991; 67: 3131-5.
94. Fitzpatrick R, Davies L. Health economics and quality of life in cancer trials: report based on a UKCCCR workshop. *Br J Cancer* 1988; 77: 1543-8.
95. Schraub S, Mercier M, Arveux P. Mesure de la qualité de vie en cancérologie. *Presse Med* 2000 ; 29 : 310-8.
96. Gotay CC, Korn EL, McCabe MS, Moore TD, Cheson BD. Quality of life assessment in cancer treatment protocols: research issues in protocol development. *J Natl Cancer Inst* 1992; 84: 575-9.
97. Rodary C, Lepage A, Hill C. Evaluation de la qualité de vie dans la recherche clinique en cancérologie. *Bull Cancer* 1998 ; 85 : 140-8.
98. Nayfield SG, Ganz PA, Moinpour CM, Cella DF, Hailey BJ. Report from a national cancer institute (USA) workshop on quality of life assessment in cancer clinical trials. *Qual life res* 1992; 1: 203-10.
99. <http://www.eortc.be/home/qol/>
100. Fossa D, Aaronson NK, Newling, et al. Quality of life and treatment of hormone resistant metastatic prostate cancer. *Eur J Cancer* 1990; 26: 1133-6.
101. Cox DR, Fitzpatrick R, Fletcher AE , et al. Quality of life assessment : we keep it simple. *J R Stat Soc, series A* 1992; 155: 353-93.
102. Zwindermann AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual life res* 1992; 1: 219-24.
103. Workshop on missing data in quality of life research in cancer clinical trials : practical and methodological issues. Guestediitors : J Bernhard and RD- Gelber. *Stat Med* 1998; 17:511-796
104. Brorsson B, Ifver j, Hays RD. The Swedish health-related quality of life survey (SWED-QUAL). *Qual life res* 1993; 2: 33-45.
105. Fairclough DL, Peterson HF, Cella D, Bonomi P. Comparison of several model-based methods for analyzing incomplete quality of life data in cancer clinical trials. *Stat Med* 1998; 17: 781-96.

106. Revicki DA, Gold K, Buckman D, Chan K, Kallich JD, Woolley M. Imputing physical function scores missing owing to mortality: results of a simulation comparing multiple techniques. *Med Care* 2001; 39: 61-71.
107. Fairclough DL. Summary measures and statistics for comparison of quality of life in a clinical trial of cancer therapy. *Stat Med* 1997; 16: 1197-209.