Valeur pronostique de la qualité de vie en cancérologie

Momar Diouf

UNIVERSITÉ DE FRANCHE-COMTÉ

ÉCOLE DOCTORALE
PRES I Bourgogne I Franche-Comté
Environnements - Santé

Ecole doctorale Environnement et Santé
Thèse de Doctorat en Sciences

Présentée et soutenue publiquement pour l'obtention du titre de
Docteur de l'université de Franche-Comté
Par Momar DIOUF
Le 17 décembre 2014

Valeur pronostique de la qualité de vie en cancérologie.

Membres du jury :

Rapporteur : Madame le Professeur Véronique Sébille, PU-PH, EA 4275 - SPHERE "Biostatistique, Pharmacoépidémiologie et Mesures Subjectives en Santé" Faculté de Pharmacie, Université de NANTES.

Rapporteur : Monsieur le Professeur Joël Coste, PU-PH, Unité de Biostatistique et d'Epidémiologie Hôtel Dieu, Paris.

Monsieur le Professeur Pascal Auquier, PU-PH, EA 3279- Laboratoire de santé publique CHU de Marseille - Hôpital de la Timone, Université de la méditerranée.

Monsieur le Docteur Stefan Michiels, PhD-HDR, Equipe de Méthodologie du service de Biostatistique et d'Épidémiologie, Institut Gustave Roussy.

Monsieur le Professeur Christophe Mariette, PU-PH, Inserm U 859, CHRU de Lille.

Monsieur le Docteur Corneel Coens, PhD, Quality of Life Department, EORTC (European Organisation for Research and Treatment of Cancer). Belgique.

Directeur de thèse : Monsieur le Professeur Franck Bonnetain, PU-PH, Responsable de l'Unité de Méthodologie et de qualité de vie en oncologie (EA3181). CHU de Besançon.

Co-Directeur de thèse : Monsieur le Docteur Thomas Filleron, PhD-HDR, Responsable Cellule Méthodologie/Biostatistique. Institut Claudius Régaud, Toulouse.

DEDICACES :

A mes parents.
A ma femme et ma fille.
A ma famille.
A mes amis.

REMERCIEMENTS :

Je souhaiterais remercier toutes les personnes qui de près ou de loin m'ont aidé à la réalisation de ce travail.

Je remercie mon directeur de thèse le Professeur Franck Bonnetain sans qui je n'aurai pas pensé faire cette thèse. Merci d'avoir accepté de diriger cette thèse avec une grande rigueur scientifique et des conseils éclairés.

Je remercie le Docteur Thomas Filleron, d'avoir accepté de codiriger cette thèse, pour son soutien, ses conseils éclairés et ses contributions originales dans ce travail.

Le Docteur Jean-Claude Barbare est plus qu'un Directeur pour moi ! Je ne le remercierai jamais assez de m'avoir motivé à faire cette thèse avec compréhension et soutien total. Je ne pouvais pas mieux tomber pour mon début dans la recherche clinique. Je te souhaite une paisible retraite.

Je remercie le réseau RICH et la DRCI du CHU d'Amiens pour leur soutien financier pour la vulgarisation de ma production scientifique.

Je remercie tous mes collègues de la Direction de la Recherche Clinique du CHU d' Amiens pour leur soutien dans ce travail.

Je remercie le Docteur Patrick Arveux et le Professeur Ziad Massy pour leurs conseils qui m'ont été d'une grande utilité.

Je remercie madame Martine Gautheron pour son aide et sa compréhension dans les conditions de réalisation de cette thèse à distance.

Je remercie toute l'équipe du Professeur Franck Bonnetain particulièrement Amélie Anota pour son aide dans ce travail.

Je remercie l'ensemble des membres de mon jury de m'avoir fait l'honneur d'accepter et pour le temps qu'ils m'ont accordé.

Je finis par remercier particulièrement ma chérie Coumba pour sa compréhension et son soutien dans ce travail. J'embrasse très fort ma fille Marianne. Désolé pour les weekends de boulot ! C'est bientôt fini ! J'espère….

Production scientifique :

Articles publiés:

1. Diouf M, Filleron T, Barbare JC, Loïc Fin, Carl Picard, Olivier Bouché, Laetitia Dahan, Xavier Paoletti, Franck Bonnetain The added value of quality of life (QoL) for prognosis of overall survival in patients with palliative hepatocellular carcinoma. *J Hepatol.* 2013; 58:509-521.
2. Diouf M, Chibaudel B, Filleron T, Tournigand C, Hug de Larauze M, Garcia-Larnicol M-L, Dumont S, Louvet C, Perez-Staub N, Hadengue A, De Gramont A, Bonnetain F. « Could baseline health-related quality of life (QoL) predict overall survival in metastatic colorectal cancer? The results of the GERCOR OPTIMOX 1 study. *Health Qual. Life Outcomes*. 2014; 12:69.

Article accepté:

1. Diouf M**,** Bonnetain F,Barbare JC, Bouché O, Dahan L, Paoletti X, Filleron T. Optimal cut-points for QLQ-C30 scales: Utility for clinical trials and updates of prognostic systems in advanced hepatocellular carcinoma. *The Oncologist.*

Article en preparation:

1. Diouf M, Filleron T, Taieb J, Bonnetain F. Prognostic value of quality of life in patients with metastatic pancreatic adenocarcinoma: a random forest methodology.

Communications orales:

1. Diouf M, Barbare JC, Filleron T, Bouché O, Dahan L, Paoletti X, Bonnetain F. Evaluation de l'apport de la qualité de vie aux scores pronostiques chez les patients atteints de carcinome hépatocellulaire en situation palliative: Résultats d'une validation externe avec l'essai CHOC.18ème journées des statisticiens des CLCC, Lille, 16-17 juin 2011.
2. Diouf M**,** Bonnetain F, Barbare JC, Bouché O, Dahan L, Paoletti X, Filleron T. Détermination d'une valeur seuil optimale pour intégrer les scores de qualité de vie du QLQ-C30 dans les scores pronostiques des patients ayant un CHC en situation palliative : comparaison de plusieurs méthodes. EPICLIN 6 / 19ème Journée des Statisticiens des Centres de Lutte Contre Le Cancer, Lyon, 9-11 mai 2012.

Communications affichées:

1. Diouf M, Bonnetain F, Chibaudel B, Tournigand C, Teixeira L, Marijon H, Perez-Staub N, De Gramont A. Could baseline health-related quality of life (QoL) improve prognostication of overall survival in metastatic colorectal cancer? Results from GERCOR OPTIMOX 1 study. J Clin Oncol 29: 2011 (suppl; abstr 3632) - ASCO Annual Meeting, Chicago (USA), 3-7 june 2011
2. Diouf M, Bonnetain F, Barbare JC, Bouché O, Meynier J, Dahan L, Paoletti X, Filleron T. Optimal cut-points for QLQ-C30 scales associated with overall survival in patients with advanced hepatocellular carcinoma (aHCC): A comparison of two methods." ESMO congress, Viennes (Autriche), 29 septembre - 2 octobre 2012.

Table des Matières

I. Liste des tableaux :

II. Liste des abréviations :

IP-OMS : Indice de performance de l'Organisation Mondial de la Santé.

Qdv : qualité de vie relative à la santé.

ACP: Adénocarcinome du pancréas.

CHC : Carcinome hépatocellulaire.

CCR: cancer colorectal métastatique.

ADN : acide désoxyribonucléique.

OMS : Organisation Mondiale de la Santé.

ASCO : American Society of Clinical Oncology.

FDA : Food and Drug Administration.

NCCN : National Comprehensive Cancer Network.

ESMO : European Society of Medical Oncology.

EASL : European Association for the Study of the Liver.

GERCOR : Groupe Coopérateur Multidisciplinaire en Oncologie.

EORTC : European Organization for Research and Treatment of Cancer.

FLIC : Functional Living Index.

FACT-G : Functional Assessment of Cancer Therapy – General.

FACT-C : Functional Assessment of Cancer Therapy - Patients with Colorectal cancer.

FACT-Hep : Functional Assessment of Cancer Therapy - Patients with Hepatobiliary cancer (liver, bile duct and pancreas).

EORTC QLQ-C30 : EORTC Quality of Life Questionnaire-Core 36.

FACIT-F : Functional Assessment of Chronic Illness Therapy-Fatigue

NRI: net reclassification improvement

IDI : integrated discrimination improvement.

CLIP : Cancer of the Liver Italian Program.

BCLC : Barcelona Clinic Liver Cancer.

GRETCH : Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire.

BoBar : BOnnetain & BARbare prognostic index.

CUPI : Chinese University Prognostic Index.

HKLC : Hong Kong Liver Cancer.

LDH : Lactate Deshydrogénase.

ASAT : Aspartate aminotransférase.

CA19-9 : antigène carbohydrate 19-9.

FOLFOX : Acide folinique/ Fluorouracile /Oxaliplatine.

FOLFIRI : Acide folinique/ Fluorouracile /Irinotecan.

FOLFIRINOX : Acide folinique/ Fluorouracile/ Irinotecan /Oxaliplatine.

III. Résumé :

Le cancer est un problème de santé publique mondial avec une estimation de 14,1 millions de nouveaux cas en 2012 pour une mortalité de 8,2 millions de personnes [1]. En France son taux d'incidence annuel est de 547/100000 habitants (355000 nouveaux cas en 2012) pour une mortalité de 148 000 personnes (source INVS).

En situation de cancer localement avancé ou métastatique, une estimation précise du pronostic est nécessaire pour un choix adéquat du traitement aussi bien dans sa nature que dans ses modalités. Cette approche fait partie du concept de la médecine de précision ou personnalisée.

L'indice de performance de l'organisation mondiale de la Santé (IP-OMS) évalué par le clinicien est souvent utilisé dans le choix du traitement. Malgré la discordance observée entre l'état général du patient évalué avec l'IP-OMS et la qualité de vie auto évaluée par le patient lui-même [2], cette dernière n'est souvent pas prise en compte dans la décision thérapeutique. Pourtant l'intérêt pronostique est établi pour plusieurs types de cancer et semble supérieur [3] ou complémentaire [4] à celui de l'IP-OMS .

Le but de ce travail est d'étudier l'apport complémentaire des scores de qualité vie relative à la santé (qdv) (par rapport à l'IP-OMS mais aussi à d'autres critères clinico-biologiques) dans l'estimation du pronostic des patients atteints de trois types de cancer, d'envisager leur intégration dans les systèmes de classification pronostique existants et de proposer des valeurs seuil qui pourraient favoriser une plus grande utilisation de ces scores de qdv en pratique clinique ainsi que dans la planification des essais cliniques.

Les données provenant de trois essais dont la qdv était un objectif secondaire ont été analysées : l'essai de phase III OPTIMOX1 sur le cancer colorectal (CCR) métastatique[5], l'essai de phase III CHOC sur le carcinome hépatocellulaire (CHC)[6] et l'essai de phase II FIRGEM sur l'adénocarcinome du pancréas (ACP)[7].

Ces trois essais étaient conçus pour des patients en situation de cancer avancé ou métastatique.

Les résultats des différentes analyses ont montré qu'indépendamment de l'IP-OMS et des autres paramètres clinico-biologiques, les scores de qdv ont un intérêt pronostique dans chacune des trois localisations cancéreuses étudiées.

Concernant le CCR métastatique, seul le score de mobilité mesuré avec l'EQ-5D est un facteur pronostique indépendant alors que les scores de mobilité et douleur/gêne permettent d'améliorer les systèmes de classification de Köhne[8] et du GERCOR[9]
Le score d'impact de l'activité quotidienne mesuré avec le QLQ-C30 est la seule composante de la qdv constaté comme facteur pronostique indépendant pour les patients atteints CHC avancé.
Toujours dans le cadre du CHC, les valeurs seuil optimales pour séparer les patients en deux groupes pronostiques homogènes sont 50, 66.66, 58.33, 66.66, 0 et 33.33 pour la santé globale, l'activité quotidienne, le bien-être physique, la fatigue, la dyspnée et la diarrhée respectivement. En utilisant ces valeurs seuil, nous avons pu montrer que ces scores de qdv permettaient d'améliorer les quatre systèmes de classification étudiés.
Pour le cancer du pancréas, les trois principaux facteurs pronostiques sont l'activité quotidienne, la fatigue et la perte d'appétit évaluées avec le QLQ-C30. Le score de santé physique est sélectionné dans le modèle final mais son impact pronostique reste marginal. Une analyse par arbre décisionnel a permis de montrer que seul le score de fatigue permettait de séparer les patients en deux groupes pronostiques avec une valeur seuil de 47.
Nos résultats pourraient permettre une évaluation plus précise du pronostic à l'aide d'informations données par le patient lui-même. Ce gain de précision dans le pronostic pourrait être utile lors du choix du meilleur type de traitement ainsi que lors de l'adaptation des doses pour les patients atteints de CCR métastatique, d'ACP métastatique et de CHC avancé. Les résultats de ce travail pourraient également être utiles dans la planification des essais cliniques ainsi que pour l'interprétation des résultats qui en sont issus.

IV. Introduction :

1) Le Cancer

Les tumeurs sont le résultat d'un processus clonal. Le clone cellulaire qui émerge acquiert progressivement des propriétés rémanentes, comme l'autonomie en termes de production de facteurs de croissance, la prolifération accrue, l'insensibilité à l'apoptose. L'acquisition de ces propriétés se fait en continu, des stades les plus précoces (simple dysplasie) à la métastase, qui constitue l'étape ultime de dissémination de la tumeur. Ces dernières années, les progrès de la génomique ont

permis de montrer l'importance des anomalies génomiques dans la progression tumorale. Parmi les anomalies les plus fréquemment identifiées, on retrouve notamment des mutations somatiques dans le génome de la quasi-totalité des cellules cancéreuse. Certaines de ces mutations jouent un rôle moteur avéré dans la carcinogenèse.

Malgré d'importantes avancées sur le plan thérapeutique, le cancer reste un défi majeur pour la médecine moderne avec près de 14,1 millions de nouveaux cas diagnostiqués en 2012 dans le monde et une mortalité annuelle associée de 8,2 millions de personnes. Les cancers les plus fréquemment diagnostiqués sont les cancers du poumon, du sein et le cancer colorectal tandis que le cancer du poumon est le plus mortel suivi par les cancers du foie et de l'estomac [1].

En France le taux d'incidence annuel du cancer est de 547/100000 habitants (355000 nouveaux cas en 2012) pour une mortalité de 148 000 personnes qui en fait la première cause de mortalité devant les maladies cardiovasculaires. Les cancers les plus fréquents en France sont celui de la prostate, du sein et le cancer colorectal alors que les plus forts taux de mortalité par cancer sont liés au cancer du poumon, suivi par les cancers colorectal et du sein.

Cette forte mortalité par cancer explique la priorité que les pouvoirs publics accordent à cette pathologie à travers la création de l'Institut National du Cancer (INCa) en 2005 et la conception de l'actuel et des deux précédents Plans cancer.

2) La Qualité de vie :

La santé a initialement été définie en 1948 par l'Organisation Mondiale de la Santé (OMS) comme : « un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité » (Actes officiels de l'Organisation Mondiale de la Santé, n°. 2, p. 100). Cependant, dès le début des années 1960 le concept de qdv a fait son apparition découlant indirectement de cette définition de la santé.

Dans sa définition de 1993, l'OMS définit la qualité de vie comme « la perception qu'a un individu de sa place dans l'existence, dans le contexte de la culture et du système de valeurs dans lequel il vit, en relation avec ses objectifs, ses attentes, ses normes et ses inquiétudes ». Il s'agit donc d'un large champ conceptuel, englobant de manière complexe la santé physique de la personne, son état psychologique, son

niveau d'indépendance, ses relations sociales, ses croyances personnelles et sa relation avec les spécificités de son environnement.

La qdv relative à la santé découle de cette définition et intègre l'impact de la maladie et du traitement. Certaines conséquences indirectes de la maladie telles que la perte d'emploi ou des difficultés financières sont également prises en compte.

Bien qu'il n'existe pas de consensus autour de la définition de la qdv, elle est généralement considérée comme un concept multidimensionnel qui inclut au minimum le bien-être physique, psychologique et social mais aussi les symptômes liés à la maladie et aux traitements.

Les premières études concernant la qdv dans le cancer ont été publiées à la fin des années 1960 / début des années 1970[10][11]. Depuis, l'intérêt de l'étude de la qdv dans le cancer n'a cessé d'augmenter et en France, a été confirmé par son intégration dans les Plans cancer pour la prise en charge des patients.

Aujourd'hui de nombreux questionnaires de qdv sont disponibles pour différentes pathologies dont le cancer qui est un problème de santé publique en France et dans le monde.

Les questionnaires les plus utilisés permettent d'évaluer des sous-dimensions de la qdv reflétant le caractère multidimensionnel de la qdv.

3) La qualité de vie dans le cancer :

Avec les nombreuses avancées thérapeutiques acquises dans leur prise en charge durant ces dernières décennies, certains types de cancer sont devenus des maladies chroniques.

Pour ces patients, la qdv est devenue un objectif secondaire majeur après la « quantité de vie ». Ainsi, à « quantité de vie » égale, un traitement qui améliore la qdv du patient devrait être privilégié. L'American Society of Clinical Oncology (ASCO) et la Food and Drug Administration (FDA) considèrent la qdv comme le critère de jugement à considérer en l'absence d'effet sur la survie globale.

Par exemple dans le cas de patients atteints de glioblastomes, Gilbert et al [12] ont montré que l'ajout du bévacizumab au schéma de Stupp [13] n'avait pas d'impact sur la survie globale mais permettait de maintenir plus longtemps une bonne qdv.

Les données de qdv sont également utilisées pour améliorer la précision de l'estimation de la survie des patients atteints de cancer à un stade localement avancé ou avancé.

Une des premières études évaluant le lien entre qdv et survie globale a été publiée en 1982 par Pater et al[14] ; elle a été suivie par beaucoup d'autres qui ont montré l'utilité de la qdv dans le pronostic de différentes localisations cancéreuses comme en attestent les méta-analyses de Montazeri [15] et Quinten [3]. Actuellement, l'IP-OMS évalué par le clinicien est utilisé en routine pour guider le choix des traitements (selon les recommandations de la National Comprehensive Cancer Network (NCCN), de l'European Society for Medical Oncology (ESMO) et de l'European Association for the Study of the Liver (EASL)) mais aussi dans les essais thérapeutiques le plus souvent comme critère d'inclusion/non inclusion.

Cependant, en situation de cancer avancé, plusieurs études ont montré la valeur pronostique des scores de qdv dans des populations homogènes de patients selon l'IP-OMS[3][16]. Ceci met en exergue la présence d'informations pronostiques complémentaires de la qdv que l'IP-OMS ne met pas en évidence confirmant ainsi la nécessité d'une auto-évaluation de la santé perçue. Les scores de qdv pourraient être un outil pour améliorer l'estimation du pronostic du patient. Ainsi à partir de cette information, une meilleure adaptation du traitement et une meilleure stratification des patients dans les essais cliniques seraient possibles.

Plusieurs outils de mesure de la qdv sont disponibles pour les patients atteints de cancer. Parmi ceux-ci on peut citer les questionnaires génériques pour toutes les pathologies comme le SF36[17], l'EQ-5D[18] et les échelles visuelles analogiques mais aussi des outils génériques beaucoup plus spécifiques du cancer que sont l'index de Spitzer[19], le FLIC[20], le FACT-G[21] et le QLQ-C30[22]. Le QLQ-C30 et le FACT-G ont des modules spécifiques pour chaque localisation cancéreuse. Pour le cancer colorectal, le FACT-C[23] contient les items du FACT-G ainsi qu'une sous-dimension spécifique alors que son module complémentaire pour le QLQ-C30 est le QLQ-CR29[24], à utiliser conjointement avec le QLQ-C30. Les modules complémentaires du QLQ-C30 sont le QLQ-HCC18[25] et QLQ-PAN26[26] pour le CHC et le cancer du pancréas respectivement alors que le FACT-Hep[27] contient une sous-dimension spécifique aux cancers hépatobiliaires en général. Le nombre d'items et de dimensions de chaque questionnaire sont décrits dans le tableau 1.

1) Valeur Pronostique de la qualité de vie :

En oncologie, la valeur pronostique de la qdv a été démontrée dans de nombreuses situations, en particulier en phase métastatique [28][29][30][31][32][33][34].

Questionnaire	Type	Année	Nombre d'items	Nombre de dimensions
SF36	Générique	1992	36	8
EQ-5D	Générique	1990	5 + VAS	5*
Spitzer Index	Spécifique Cancer	1981	5	1
FLIC	Spécifique Cancer	1984	22	6+
FACT-G	Spécifique Cancer	1993	27	4
FACT-C	Spécifique CCR	1999	9**	1
FACT-Hep	Spécifique CHC et CP	2002	18**	1
EORTC QLQ-30	Spécifique Cancer	1993	30	15
EORTC QLQ-CR29	Spécifique CCR	2009	29	17
EORTC QLQ-HCC18	Spécifique CHC	2004	18	8
EORTC QLQ-PAN26	Spécifique CP	1999	26	10

*contient également une échelle visuelle analogique (EVA) pour évaluer la santé globale.
+Seul questionnaire constitué essentiellement d'EVA, les autres étant sous forme d'échelle de Likert.
**Uniquement la partie spécifique à la localisation car FACT-Hep et FACT-C contiennent le FACT-G par définition.
CCR=cancer colorectal; CHC=carcinome hépatocellulaire ; CP=cancer du pancréas.
Tableau 1 : Caractéristiques de quelques questionnaires parmi les plus utilisés.

La dimension de santé physique et les symptômes « douleur » et « perte d'appétit » étaient notamment corrélés à la survie globale dans une méta-analyse sur données individuelles regroupant onze pathologies cancéreuses [3]. Dans cette étude, l'IP-OMS n'était pas significativement associé à la survie en présence des trois dimensions de qdv ci-dessus. La santé physique auto-évaluée par le patient semble contenir plus d'information pronostique que l'IP-OMS qui résume l'état général du patient évalué par le clinicien. D'autre part, la valeur des symptômes « douleur » et « perte d'appétit » est très intéressante ; leur utilisation pourrait notamment améliorer la communication patient/médecin ainsi que le choix de la stratégie thérapeutique.
La communication pourrait s'améliorer dans le sens où le clinicien ne pouvait probablement pas avoir connaissance de tels symptômes sans l'évaluation de la qdv, ce qui lui permet d'en tenir compte et le patient à son tour sentant son avis important pourrait s'impliquer plus et poser davantage de questions.

Type de cancer	Stade	Auteur	Année	Questionnaire	Résultats
Poumon non à petites cellules au	non opérable	Langendijk	2000	EORTC QLQ-C30	Score de santé globale
Poumon non à petites cellules	Avancé	Moinpour et al.	2002	FACT-L	Score de global du FACT-L
Poumon non à petites cellules	Avancé	Efficace	2006	EORTC QLQ-C30+ EORTC QLQ-LC13	Douleur et dysphagie
Poumon non à petites cellules	Hétérogène nouvellement diagnostiqués	Jacot	2008	LCSS	Score global du LCSS
Sein	Avancé	Coates	1992	Echelle visuelle analogique pour santé physique+humeur, nausée, vomissement et perte d'appétit.	L'index de qdv et de score de santé physique.
Sein	Avancé	Luoma	2003	EORTC QLQ-C30	Douleur
Sein	métastatique	Efficace	2004	EORTC QLQ-C30 + QLQ-BR23	Perte d'appétit
Cancer de l'œsophage	Avancé	Park	2008	EORTC QLQ-C30	Score de bien-être social
Cancer oesophaso-gastrique.	Localement avancé et métastatique	Yau	2004	EORTC QLQ-C30	Activité quotidienne, bien-être physique et score de santé globale.
Estomac		McKernan	2008	EORTC QLQ-C30	Perte d'appétit
Rein		Cella	2012	FKSI+ FKSI-DRS+FACT-G	Les trois scores globlaux du FKSI, du FKSI-DRS et du FACT-G.
Tête et cou	Localement avancé	Coyne	2007	FACT-G	Bien-être émotionnel
Myélome multiple	Nouvellement diagnostiqués	Wisloff and Hjorth	1997	EORTC QLQ-C30	Bien-être physique
Cerveau	Métastatique et non métastatique	Sehlen et al.	2003	FACT-G	Score global du FACT-G
Glioblastome	Grade IV	Mauer	2007	EORTC QLQ-C30 + QLQ-BN20	Bien-être social et cognitif ainsi le score de santé global
Vessie	Métastatique	Roychowdury	2003	EORTC QLQ-C30	Activité quotidienne, bien-être physique et anorexie

EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
EORTC QLQ-BN20: module complémentaire du QLQ-C30 spécifique aux patients atteints de cancer du cerveau.
EORTC QLQ-BR23: module complémentaire du QLQ-C30 spécifique aux patients atteints de cancer du sein.
EORTC QLQ-LC13: module complémentaire du QLQ-C30 spécifique aux patients atteints de cencer du poumon.
FACT-G: Functional Assessment of Cancer Therapy - General
FACT-L: Functional Assessment of Cancer Therapy - Lung
FKSI: Functional Assessment of Cancer Therapy–Kidney Symptom Index
FKSI-DRS: sous échelle du FKSI pour évaluer les symptômes liés au cancer.
LCSS=Lung Cancer Symptoms Scale.
Tableau 2 : Valeur pronostique de la qualité de vie en fonction du type de cancer.

Cependant, une approche globale de la qdv telle que présentée dans l'étude de Quinten et al. [3] reste problématique étant donnée l'hétérogénéité du pronostic en fonction de la localisation cancéreuse et du stade de la maladie. En effet, on ne peut

pas exclure que la contribution de chaque score de qdv soit différente en fonction du site du cancer primitif et donc qu'étudier la relation entre survie et qdv pour chaque localisation cancéreuse soit plus pertinent.

Dans une autre étude, Quinten et al[35] ont évalué la valeur pronostique des scores QLQ-C30 pour la survie globale en fonction de la pathologie cancéreuse. Plusieurs publications ont synthétisé pour chaque localisation cancéreuse les dimensions de qdv pertinentes pour le pronostic (Tableau 2).

L'hétérogénéité des dimensions retrouvées à travers les différentes localisations pourrait en partie s'expliquer par la différence d'outils utilisés en plus d'une probable différence de contribution des dimensions de qdv sur le pronostic des différents types de cancer. La différence de méthodologie utilisée dans les différentes études pourrait également contribuer aux différences observées.

L'utilisation d'un outil unique pour les études pronostiques (par exemple le QLQ-C30) et d'une méthodologie statistique standardisée pourrait permettre de trancher entre les différentes possibilités. Pour éviter le phénomène de multicolinéarité, les scores de qdv avec une taille d'effet clinique non négligeable et dont la valeur pronostique est cohérente avec le mécanisme biologique du cancer étudié devraient être privilégiés pour entrer dans le modèle multivarié.

Le tableau 3 résume les résultats de certaines études parmi celles qui se sont intéressées à l'intérêt pronostique de la qdv dans les trois localisations cancéreuses étudiées.

i. Cancer colorectal métastatique :

Le cancer colorectal est le troisième cancer le plus diagnostiqué au monde et le quatrième cancer de plus mauvais pronostic [1]. Au moment du diagnostic, 35% des patients ont des métastases[36] alors qu'entre 20 et 50% des patients sans métastases en développeront plus tard [36][37]. En situation métastatique, la médiane de survie varie entre 6 et 30 mois [38][39] selon le type de chimiothérapie et le statut mutationnel du gène KRAS.

Plusieurs études se sont intéressées à évaluer la valeur pronostique de la qdv dans le cancer colorectal métastatique[40][41][42].

L'étude de Maisey et al. [42] utilisant le QLQ-C30 a montré que le scores de santé globale, de fonction physique, sociale, émotionnelle et d'activité quotidienne ainsi que les symptômes de douleur, nausée, dyspnée et insomnie étaient des facteurs

pronostiques indépendants. Cependant la méthodologie utilisée ne valide pas le caractère « indépendant » de chacun de ces différents scores par rapport aux paramètres clinico-biologiques. L'intérêt propre de chaque score de qdv par rapport aux autres scores n'était également pas étudié. Chaque score de qdv était inclus dans un modèle construit avec des variables clinico-biologiques (modélisation pas à pas) alors qu'un modèle incluant tous les scores en plus des paramètres clinico-biologiques aurait permis de confronter l'utilité des différents scores de qdv entre eux même si le phénomène de multicolinéarité est une faiblesse pour ce dernier modèle. Il est fort probable que l'on n'ait pas autant de scores de qdv reconnus comme des facteurs pronostiques indépendants s'ils étaient tous inclus dans un même modèle multivarié. Une autre approche à privilégier aurait été de faire un travail de sélection préliminaire des scores de qdv en les regroupant en classes de variables corrélées et de choisir un représentant pour chaque classe tout en privilégiant la facilité d'interprétation du représentant de chaque classe.

Braun et al.[40] ont montré que la perte d'appétit était un facteur indépendant de survie.

En utilisant une méthodologie plus rigoureuse, Efficace et al. [41] ont montré en 2006 que seule la dimension de bien-être social était un facteur pronostique indépendant de survie après ajustement sur 3 facteurs clinico-biologiques validés. En 2008, avec une seconde cohorte le même auteur a validé la valeur pronostique de la dimension sociale en utilisant le modèle initialement développé en 2006. Pour les autres dimensions de la qdv, seuls les résultats de leur impact pronostique en analyse univariée de Cox ont été donnés [43].

ii. Carcinome hépatocellulaire :

Le cancer du foie dont le carcinome hépatocellulaire (CHC) représente 70 à 85% des cas est le 5ème cancer le plus diagnostiqué et la 2ème cause de mortalité par cancer dans le monde [1]. Au moment du diagnostic, à peu près 70% des patients ont une maladie considérée comme incurable [44].

Contrairement aux autres localisations tumorales, la valeur pronostique de la qdv a peu été étudiée dans le CHC [45][46].

Une étude réalisée en Chine a utilisé le QLQ-C30 dans une population de patients atteints de CHC d'étiologie virale B. Dans cette étude, la perte d'appétit, la santé physique et le score d'activité quotidienne étaient des facteurs pronostiques indépendants. L'hépatopathie sous-jacente étant très importante dans le pronostic

des patients atteints de CHC, ces résultats ne sont pas directement transposables aux malades européens dont l'hépatopathie est majoritairement d'origine alcoolique ou due au virus de l'hépatite C.

Etude	Outil	Méthode	Résultats	Avantages	inconvénients
Cancer colorectal					
Maisey N=497 Nq=299 D=253	QLQ-C30	MCPP HP+ MC+	Dimensions physique, activité quotidienne, sociale et émotionnelle, santé globale, douleur, nausée, dyspnée et insomnie	Méthodologie intéressante évitant le problème de multicolinéarité	Ne facilite pas l'utilisation de la qdv en routine car pas beaucoup de dimensions pronostiques et absence de modèle multivariée incluant toutes les dimensions. Ni validation interne, ni externe.
Braun N=396 Nq=396 D=211	QLQ-C30	MCSS HP+ MC+	Perte d'appétit.	Méthodologie rigoureuse pour l'analyse pronostique des données de qdv. Validation interne	Pas de validation externe.
Efficace Apprentissage : N=497 Nq=299 D=253 Validation N=564 Nq=443 D=354	QLQ-C30	MCS HP+ MC+	Dimension sociale.	Méthodologie rigoureuse. Validation interne et externe	Faible différence de C-index entre le modèle avec et sans données de qdv (0.629 à 0.648). La significativité de cette différence n'a pas été étudiée.
Carcinome hépatocellulaire					
Yeo N=233 Nq=233 D=209	QLQ-C30	MCS HP- MC-	Perte d'appétit, dimensions physique et activité quotidienne.	Première étude sur qdv pronostique dans le CHC avec le QLQ-C30	Ni validation interne, ni externe.
Bonnetain	Index de	MCS	Qualité de vie globale	Méthodologie	Pas de

N=538 Nq=489 D=459	Spitzer	HP- MC*		rigoureuse. Validation interne.	validation externe même si elle est prévue dans l'article.
			Cancer du pancréas		
Lis	Index de qualité de vie de Ferrans et Powers		« Santé globale et physique »	Méthodologie rigoureuse.	Outil générique n'évalue pas les symptômes liés au cancer. Ni validation interne, ni externe.
Robinson N=86 Nq=86 D=NA	FACIT-F SF-36 FAACT BPI	MCS HP- MC-	Fatigue	Facteurs d'ajustement prédéfinis.	Nombreux questionnaires utilisés et méthodologie statistique pas suffisamment détaillée. Le choix des facteurs d'ajustement n'est pas argumenté. Absence d'indices de performances des modèles.
Bernhard N=311 Nq=299 D=NA	Echelle visuelle analogique	MCSS HP- MC-	Douleur et fatigue	Absence de méthode stepwise et dimensions de qdv choisies en fonction de la connaissance du sujet.	Ni validation interne, ni externe.
Gourgou N=342 Nq=320 D=273	QLQ-C30	MCS HP+ MC-	santé physique, constipation, dyspnée.	Méthodologie rigoureuse.	Ni validation interne, ni externe.

Braun N=186 Nq=186 D=NA	QLQ-C30	Santé globale	Méthodologie rigoureuse pour l'analyse pronostique des données de qdv. Validation interne	Pas de validation externe.

MCPP= Modèle de Cox pas à pas.
MCSS=Modèle de Cox sans sélection stepwise.
MCS=Modèle de Cox avec sélection stepwise.
MC*= Multicolinéarité non aplicable
MC+=Prise en compte Multicolinéarité
MC-=Absence de prise en compte Multicolinéarité
HP+=Hypothèse de hazard proportionnel vérifiée.
HP-=Hypothèse de hazard proportionnel non vérifiée.
FACIT-F=Functional Assessment of Chronic Illness Therapy–Fatigue
FAACT =Functional Assessment of Anorexia/Cachexia Therapy
BPI=Brief Pain Inventory
N=effectif total de l'étude.
Nq=Nombre de ptaients ayant complété les données de qualité de vie.
D=Nombre de décès.
NA=Non renseigné.

Tableau 3 : Récapitulatif de l'analyse pronostique de la qualité de vie dans les trois localisations de cancer étudiées (colon-rectum, foie et pancréas).

Pour les patients atteints de CHC d'étiologie majoritairement alcoolique, Bonnetain et al.[46] ont montré que la qdv évaluée par l'index de Spitzer était un facteur pronostique indépendant. Même si cet index ne permet pas de faire une analyse plus fine des différentes composantes de la qdv, il a contribué de façon significative à l'amélioration de la performance des quatre systèmes de classification étudiés : CLIP, BCLC, GRETCH et BoBar.

i. Cancer du pancréas :

Avec 338000 nouveaux cas diagnostiqués chaque année dans le monde, le cancer du pancréas est le douzième cancer le plus diagnostiqué au monde et le septième plus mortel [1] dont l'adénocarcinome du pancréas constitue la grande majorité (entre 80 et 90%).

Au moment du diagnostic, 80% des patients sont en situation métastatique avec une médiane de survie allant de 6 mois (sous Gemcitabine) à 11 mois (sous FOLFIRINOX)[47]

Lis et al[48] ont retrouvé la dimension « santé et activité physique » comme facteur indépendant de survie en utilisant l'index de qualité de vie de Ferrans and Powers [49]. Cependant cet index n'est pas couramment utilisé dans les études cliniques dans le cancer, relativisant la pertinence des résultats.

Robinson et al.[50] ont retrouvé la fatigue comme seul composant de la qdv indépendamment associé à la survie des patients à l'aide du questionnaire FACIT-F.
Une étude plus récente de Bernhard et al. [51] confirme le caractère pronostique de la fatigue, en plus de la douleur, à l'aide d'une échelle visuelle analogique.
Dans un essai thérapeutique comparant la Gemcitabine au FOLFORINOX, Gourgou et al.[52] ont montré que la dimension de santé physique, la constipation, la dyspnée évaluées à l'aide du QLQ-C30 étaient des facteurs indépendants de survie. Même si elle est sélectionnée dans un des modèles de cette étude, le rôle pronostique de la douleur n'est pas clairement établi.
Enfin, Braun et al. [53] ont montré que le score de santé globale est un facteur indépendant de survie.

5) Système de classification pronostique:

i. Introduction :

Les systèmes de classification pronostique sont essentiels en médecine, et notamment en oncologie. En effet, le choix d'un traitement pour un patient dépend du rapport bénéfice/risque associé à ce traitement, tandis que ce rapport bénéfice/risque dépend à son tour du pronostic du patient. Une estimation la plus précise possible du pronostic est donc indispensable pour informer le patient et pour prendre la décision thérapeutique. Les systèmes pronostiques sont également utiles pour définir l'éligibilité des patients dans les essais cliniques (la balance bénéfice/risque pouvant être défavorable à un patient de bon pronostic) et pour la stratification de la randomisation surtout dans le cas d'un effectif limité.
Pour construire un système de classification pronostique, l'analyste essaie de construire un modèle qui reflète au mieux les données observées. Cependant, un manque de rigueur dans certains aspects comme la sélection des variables candidates, la vérification des hypothèses du modèle utilisé et la manière dont les données manquantes ont été traitées peuvent aboutir à la construction d'un modèle de faible performance lors de sa validation externe.

ii. Construction :

Avant la construction d'un modèle pronostique, la variable à expliquer et les variables explicatives potentielles doivent être identifiées. Les variables explicatives potentielles devraient être choisies à l'aide d'un travail important de « nettoyage » en

supprimant les variables dont l'absence d'information pronostique a clairement été établie par des études antérieures et en sélectionnant d'emblée celles dont la valeur pronostique est partiellement ou totalement établie. Concernant les variables non encore étudiées, la significativité statistique devrait être accompagnée d'une taille d'effet clinique non négligeable pour qu'elles soient candidates pour le modèle multivarié.

Afin d'évaluer le pronostic des patients atteints de cancer, les nomogrammes ont été longtemps utilisés.

A l'heure actuelle, les scores pronostiques sont construits à l'aide d'un modèle statistique qui dépend de la nature des données à modéliser. Dans le cadre du cancer, les déterminants de la durée de vie des patients ou de façon équivalente du risque instantané de décès sont recherchés avec le modèle semi-paramétrique de Cox[54] le plus souvent. Des méthodes paramétriques alternatives dont celle de Weibull et de Gompertz[55] existent mais nécessitent de vérifier des hypothèses plus contraignantes que celles du modèle de Cox; ce qui explique que ces méthodes paramétriques soient moins utilisées que celui de Cox.

Les arbres de classification constituent une alternative pour le développement de modèles pronostiques mais requièrent un choix adéquat de la profondeur de l'arbre pour un équilibre entre précision et stabilité. La profondeur de l'arbre est définie comme la distance qui sépare la dernière variable utilisée pour diviser un sous-groupe de patients de la racine de l'arbre, qui correspond à la première variable permettant de diviser les patients en deux groupes les plus homogènes possibles. En d'autres termes si l'on appelle nœud tout endroit de partitionnement d'une variable, la profondeur d'un arbre est la longueur du trajet entre le nœud initial (racine) et le nœud terminal (feuille).

Plus récemment, la méthodologie des forêts aléatoires a prouvé sa supériorité par rapport aux arbres de classification et parfois au modèle de Cox pour la recherche des facteurs pronostiques mais l'utilisation des résultats de ces forêts aléatoires pour construire un modèle pronostique n'est pas encore très élaborée[56].

Cependant, quelle que soit la stratégie de développement d'un modèle, les étapes suivantes devraient être respectées :

- Choix des variables à étudier en essayant de respecter la règle de 10 évènements pour une variable [57] afin de minimiser l'inflation du risque

de première espèce. Cependant une récente étude [58] a montré que cette règle pourrait être assouplie jusqu'à 5 évènements par variable.

- Construction d'un modèle multivarié soit avec les méthodes de sélection « stepwise », soit avec les méthodes de maximum de vraisemblance pénalisé ou bien avec construction raisonnée sans stepwise à partir de variables présélectionnées.
- Evaluation de la monotonicité du gradient pour montrer que les patients dans un groupe de moins bon pronostic vivent moins longtemps que ceux d'un groupe de bon pronostic.
- Vérifications a posteriori des hypothèses du modèle (log-linéarité et hasard proportionnel pour le modèle de Cox).
- Mesure de calibration avec par exemple des tests de Hosmer-Lemeshow à des instants prédéfinis.
- Mesure des indices de performance du modèle : le C-index d'Harrell[55], la statistique de Schemper[59], le NRI (Net Reclassification Improvement) et l'IDI (Integrated Discrimation Improvement) de Pencina [60]. Ces paramètres sont définis dans le tableau 4.
- Analyse de sensibilité par rapport aux données manquantes[61][62] si elles sont fréquentes.

Pour le cas particulier de l'évaluation de l'intérêt pronostique de la qdv, il faudrait une comparaison des performances entre le modèle avec les paramètres clinico-biologiques et démographiques (modèle sans données de qdv) et le modèle sans données de qdv ci-dessous auxquels les scores de qdv ont été ajoutés (modèle avec qdv tout en forçant les paramètres traditionnels dans le modèle) [63].

i. Validation :

Tout modèle pronostique devrait faire l'objet d'une validation externe avec une seconde cohorte indépendante (cohorte de validation) de celle ayant servi à sa construction (cohorte d'apprentissage). Si l'effectif le permet, la cohorte totale peut être divisée en deux échantillons : un échantillon d'apprentissage et un échantillon de validation avec généralement une répartition 2/3 et 1/3.

Une validation interne est un préalable nécessaire pour une validation externe car il serait très peu probable qu'un modèle sans validité interne ait une validité externe.

Étant donné que la majorité des modèles développés n'ont pas de validation externe, une validation interne devrait être réalisée pour les études avec un effectif modéré en attendant une éventuelle validation externe[64]. Parmi les méthodes de validation interne on peut citer la validation croisée[65] et le ré-échantillonnage bootstrap[55] qui permettent de calculer des indices de performance corrigés de l'optimisme du modèle initial.

Formules	Interprétation	Avantages	inconvénients
C-index **C-index=P[$T_i > T_j$ / $S(t/X_i) > S(t/X_j)$]** **-T** est la survie observée et **S(t/X)** est la survie prédite sachant la covariable **X** -Sachant que les paires i et j sont comparables. Deux paires ne sont pas comparables si celui dont le suivi est plus petit correspond à une censure.	-Indice de discrimination du modèle. -Varie entre 0.5 (absence de discrimination) et 1 (parfaite discrimination)	Mesure robuste de performance.	-Pas intuitif cliniquement en termes d'interprétation.
Statistique de Schemper. $V = \frac{D - D_x}{D}$ **-D** est la précision du modèle sans les variables étudiées. **-Dx** est la précision du modèle avec les variables étudiées.	-Proportion de variance expliquée similaire au R^2 pour le modèle linéaire. -Gain de précision dans l'estimation de la survie.	Statistiquement robuste.	-Pas intuitif cliniquement en termes d'interprétation.
NRI^2 **NRI= P(Up/Event) – P(Down/Event) + P(Down/Non-Event) – P(Up/Non_Event)** **-« Up »** signifie que la probabilité prédite par le nouveau modèle est supérieure à celle de l'ancien modèle. **-« Down»** signifie que la probabilité prédite par le nouveau modèle est inférieure à celle de l'ancien modèle. **-Event**= Patient ayant l'évènement attendu. **-Non_Event**= Patient n'ayant pas l'évènement attendu à l'instant choisi.	Indice de gain en reclassification des patients selon leur risque.	-Quantification du gain de reclassifications correctes. -Facilité d'interprétation.	-N'est pas une proportion [66] (valeur maximale égale à 2). -Valeur minimale importante non définie. -Sensible au cutoff pour définir les groupes de risque. -Ne prend pas en compte la prévalence de l'évènement ; donc sans pondération peut aboutir à une fausse conclusion d'un meilleur modèle quand celui-ci fait pire que l'ancien.

½ NRI (>0) : ie NRI sans catégories **$P(Q_i > P_i /Event) + P(Q_i < P_i /Non\_Event)$** **$-Q_i$** est la probabilité d'évènement pour le sujet i d'après le nouveau modèle. **$-P_i$** est la probabilité d'évènement pour le sujet i d'après l'ancien modèle.	Version continue du NRI.	Facilité d'interprétation.	-N'est pas une proportion. -Ne mesure pas forcément l'utilité d'un nouveau marqueur (Si ancien modèle mal calibré). -Ne mesure pas la différence de performance entre un ancien et un nouveau modèle mais une différence intra-individuelle de probabilité. L'amplitude de la différence entre Q_i et P_i n'est pas prise en compte. -On peut un « NRI>0 » élevé sans gain en discrimination et donc absence d'impact sur la prise en charge.
IDI : $\int_0^1 \mathbf{sens(u)du} + \int_0^1 \mathbf{(1 - sep(u))du}$ *Sens(u)* et *sep(u)* représentent respectivement la sensibilité et la spécificité correspondant à la valeur seuil u	Différence de pente de discrimination (discrimination slope) entre les 2 modèles. Pente de discrimination=différence de probabilité d'évènement chez ceux qui en ont et ceux qui n'en ont pas.	-Permet d'évaluer l'amplitude du gain lié à l'ajout de la nouvelle variable. -Facilité d'interprétation	Absence de valeur minimale pour juger l'importance du gain de discrimination.

NRI=Net Reclassification Improvement.
NRI^2 =NRI pour deux classes.
NRI(>0)=NRI continue (sans catégories).
IDI=Integrated Discriminant Improvement.

Tableau 4 : Définition et interprétation des paramètres de performance d'un modèle pronostique.

Une forte variabilité de ces indices pourrait laisser penser à une instabilité du modèle et donc à une faible validité interne. En plus de ces différentes étapes dans la construction d'un modèle pronostique, le problème de la multicolinéarité est une difficulté supplémentaire pour les données de qdv dans le modèle multivarié. Ainsi pour les données de qdv, Mauer et al.[63] recommandent le calcul de la matrice des corrélations ou des facteurs d'inflation de la variance pour identifier les groupes de variables fortement corrélées et d'en choisir un représentant pour chaque groupe. Une comparaison de la survie globale et des données clinico-biologiques entre les patients avec et sans données de qdv devrait être réalisée pour comprendre le

mécanisme des données manquantes. Si l'hypothèse d'une cause complètement aléatoire n'est pas plausible, une imputation des données manquantes et une analyse de sensibilité des résultats par rapport aux données manquantes devraient être systématiquement réalisée. Mauer et al[63] insistent également sur la nécessité de mener plus de validations externes des modèles étudiant l'intérêt pronostique de la qdv dans le cancer.

ii. Scores pronostiques existants :

Plusieurs scores pronostiques existent pour le cancer colorectal métastatique dont le système de classification de Glasgow [67], le sous-système de classification pour le stade IV de l'AJCC (American Joint Committee on Cancer)[68], le score de Köhne [8] et le système pronostique du GERCOR [9].

Plusieurs systèmes de classification ont été développés pour le carcinome hépatocellulaire, les principaux étant dénommés Okuda [69], CLIP[70], BCLC[71], GRETCH, BoBar[72], Glasgow[73], CUPI[74], et JIS [75]. Le système BCLC reste à ce jour le plus utilisé car il est associé à un algorithme décisionnel pour le traitement. Récemment, un nouveau système de classification HKLC[76] a été développé à partir de 3856 patients asiatiques atteints de CHC.

Pour le cancer du pancréas métastatique, un système de classification sous forme de nomogramme a été développé par Hamada et al.[77]. Vernerey et al. ont développé un nomogramme pour les patients atteints de cancer du pancréas localement avancé (http://www.fondationarcad.org/nos-actions/programmes-de-recherche/Prognostic-Score-Nomogram-OS-in-LAPC). Ce nomogramme pourrait être testé chez les patients atteints de cancer du pancréas métastatique.

6) Score pronostique et qualité de vie :

A notre connaissance, aucun système de classification pronostique intégrant les données de qdv pour les patients atteints de cancer n'est proposée dans la littérature médicale malgré la démonstration de la valeur ajoutée de ces scores de qdv par rapport à l'IP-OMS pour l'évaluation du pronostic dans le cancer en général [3] et en particulier dans le CHC [46][45], le cancer colorectal[41] et l'adénocarcinome du pancréas[52]. Une récente étude de Hsu et al. [78] a montré qu'une réaffectation des patients avec IP-OMS 0-1 du stade avancé au stade intermédiaire permet d'améliorer le système de classification original de Barcelone (BCLC). Il parait donc très probable que le manque de valeurs seuil rigoureusement établies et

unanimement acceptées limite l'utilisation pratique des scores de qdv aussi bien en routine clinique que dans les essais thérapeutiques.

Dans leur récente analyse critique des causes probables d'échec des essais de phase III dans le CHC avancé, Llovet et al. [79] suggèrent qu'un excès de mortalité lié à la toxicité du traitement dans le bras expérimental pourrait en être une explication. Cela pourrait s'expliquer par une faible sensibilité des scores pronostiques existants pour identifier les patients fragiles et vulnérables aux effets secondaires des médicaments étudiés. L'auto-évaluation de la qdv du patient pourrait ainsi permettre d'identifier de façon plus précise ces patients qui devraient être exclus des futurs essais cliniques.

Une auto-évaluation par le patient de son état de santé pourrait donc aider le clinicien à ajuster le type de traitement ainsi que la dose acceptable pour éviter une surmortalité liée au traitement plutôt qu'au cancer lui-même.

La pratique médicale moderne tendant de plus en plus vers une prise en décision thérapeutique partagée entre le médecin et son patient, une telle absence des scores de qdv dans les systèmes pronostiques constitue un frein à l'utilisation des données de qdv dans la prise en charge des patients.

Peu d'études se sont intéressées à la valeur prédictive de la qdv pour un autre évènement en dehors du décès. Parmi ces études, on peut citer celle de Siddiqui et al.[80] qui a montré l'intérêt du score de bien-être physique pour prédire le contrôle locorégional après radiothérapie pour les patients atteints de cancer de la tête et du cou en situation localement avancée. Sarenmalm et al.[81] ont montré que les scores de bien-être physique et de nausée/vomissement permettaient de prédire la rechute des patients atteints de cancer du sein et recevant une chimiothérapie adjuvante.

7) Problématique de la thèse :

Une évaluation précise du pronostic est un préalable indispensable pour une optimisation de la prise en charge des patients atteints de cancer. Or plusieurs études ont montré que la performance des systèmes de classification existants était très perfectibles.

Le but de ce travail a été d'étudier l'apport complémentaire des scores de qdv (par rapport à l'IP-OMS mais aussi à d'autres critères clinico-biologiques) dans l'estimation du pronostic des patients atteints de trois types de cancer, d'envisager leur intégration dans les scores pronostiques existants et de proposer des valeurs

seuil qui pourraient favoriser une plus grande utilisation de ces scores de qdv en pratique clinique ainsi que dans la planification des essais cliniques.

Ainsi, ***dans la première partie du travail*** dont les résultats sont synthétisés dans le 1er article, nous chercherons à montrer l'intérêt pronostique de la qdv dans le cancer colorectal métastatique en utilisant le questionnaire EUROQOL EQ-5D. Une mise à jour (incluant les scores de qdv) de deux des principaux systèmes de classification pronostique les plus connus sera proposée en utilisant les données de 620 patients issus de l'essai de phase III OPTIMOX1 promu par le groupe GERCOR dont l'analyse de la qdv était un objectif secondaire. L'essai OPTIMOX1 avait pour objectif de montrer la supériorité d'une administration séquentielle de FOLFOX par rapport à une administration continue de FOLFOX.

Dans la deuxième partie portant sur le carcinome hépatocellulaire, le 2ème article résume le travail de validation de la valeur pronostique de la qdv évaluée avec le questionnaire QLQ-C30 [22] de l'EORTC. Les résultats de la recherche de valeurs seuil pour les scores de qdv prédictifs de la survie globale et de révision de quatre systèmes de classification pronostique (intégrant les scores de qdv ainsi dichotomisés) figurent dans le 3ème article.

Les données de 271 patients provenant de l'essai CHOC dont l'objectif principal était de montrer l'efficacité de l'octréotide-retard dans le traitement du CHC en situation avancée sont analysées.

Dans la troisième partie concernant l'adénocarcinome du pancréas, l'intérêt pronostique de la qdv sera étudié et un système de classification pronostique établi sous forme d'arbre décisionnel construit à l'aide de variables dont la valeur pronostique a été préalablement validée avec une méthodologie de forêts aléatoires appliquées aux données de survie.

Dans cette partie, les données de 98 patients de l'essai de phase II FIRGEM dont le promoteur est l'Association des Gastro-Entérologues Oncologues (AGEO) seront utilisées. La qdv était évaluée avec le questionnaire QLQ-C30 [22] de l'EORTC.

L'essai FIRGEM avait pour objectif principal de montrer la supériorité d'un traitement par Gemcitabine associé à une administration séquentielle de FOLFIRI par rapport à un traitement par Gemcitabine seul pour les patients atteints d'adénocarcinome du pancréas métastatique non prétraité. Le 4ème article résume le travail effectué dans cette dernière partie.

IV. Articles :

i. Qualité de vie et cancer colorectal métastatique :

1) Résumé :

Rationnel: La valeur pronostique de la qdv a été étudiée dans plusieurs types de cancer. Une récente étude a montré que les systèmes de classification pronostique du cancer colorectal (CCR) métastatique sont améliorables. Nous avons évalué l'intérêt pronostique de la qdv dans le cancer CCR métastatique et sa contribution à l'amélioration des performances des systèmes de classification de Köhne et du GERCOR.

Méthode: Le questionnaire EUROQOL EQ-5D était complété par les patients avant la randomisation dans l'étude de phase III OPTIMOX1 dont l'objectif principal était de comparer deux stratégies de chimiothérapie par FOLFOX. 620 patients atteints de CCR métastatique initialement non traités ont été inclus dans cette étude entre janvier 2000 et juin 2002 à travers 56 centres dans cinq pays. L'amélioration des performances des systèmes de classification (après ajout des scores de qdv) a été étudiée avec l'indice de discrimination de Harrell et l'indice NRI.

Résultats: Parmi les 620 patients, 249 (40%) ont complété les données de qdv**.** Le système de Köhne a pu être amélioré par le LDH, la mobilité et la douleur/gêne; l'indice de Harrell a augmenté de 0.54 à 0.67. Le NRI à 12 mois était de 0.23 (IC95%=[0.05; 0.46]). La mobilité et la douleur/gêne ont pu améliorer le système du GERCOR: l'indice de Harrell a augmenté de 0.63 à 0.68 et l'indice NRI à 12 mois était de 0.35 [0.12; 0.44].

Conclusion: Les dimensions de mobilité et de douleur/gêne étaient des facteurs pronostiques indépendants et pourraient être utiles pour la classification et le choix du traitement pour les patients atteints de CCR métastatique.

2) Article sur le cancer colorectal:

Diouf et al. Health and Quality of Life Outcomes 2014, 12:69
http://www.hqlo.com/content/12/1/69

HEALTH AND QUALITY OF LIFE OUTCOMES

RESEARCH Open Access

Could baseline health-related quality of life (QoL) predict overall survival in metastatic colorectal cancer? The results of the GERCOR OPTIMOX 1 study

Momar Diouf[1,2*], Benoist Chibaudel[3], Thomas Filleron[5], Christophe Tournigand[6], Marine Hug de Larauze[4], Marie-Line Garcia-Larnicol[3], Sarah Dumont[3], Christophe Louvet[7], Nathalie Perez-Staub[3], Alexandra Hadengue[4], Aimery de Gramont[3] and Franck Bonnetain[2,4]

Abstract

Background: Health-related quality of life (QoL) has prognostic value in many cancers. A recent study found that the performance of prognostic systems for metastatic colorectal cancer (mCRC) were improvable. We evaluated the independent prognostic value of QoL for overall survival (OS) and its ability to improve two prognostic systems'performance (Köhne and GERCOR models) for patients with mCRC.

Methods: The EQ-5D questionnaire was self-completed before randomization in the OPTIMOX1, a phase III trial comparing two strategies of FOLFOX chemotherapy which included 620 previously untreated mCRC patients recruited from January 2000 to June 2002 from 56 institutions in five countries. The improvement in models' performance (after addition of QoL) was studied with Harrell's C-index and the net reclassification improvement.

Results: Of the 620 patients, 249 (40%) completed QoL datasets. The Köhne model could be improved by LDH, mobility and pain/discomfort; the C-index rose from 0.54 to 0.67. The associated NRI for 12-month death was 0.23 [0.05; 0.46]. Mobility and pain/discomfort could be added to the GERCOR model: the C-index varied from 0.63 to 0.68. The NRI for 12 months death was 0.35 [0.12; 0.44].

Conclusions: Mobility and pain dimensions of EQ5D are independent prognostic factors and could be useful for staging and treatment assignment of mCRC patients. Presented at the 2011 ASCO Annual Meeting (#3632).

Background

Colorectal cancer (CRC) is the third most diagnosed cancer in men and the second most diagnosed in women, with over 1.2 million new cases and 608 700 deaths worldwide in 2008 [1]. About up to half (20% to 50%) of CRC patients will develop metastases during the course of their disease [2] and approximately 35% are diagnosed with synchronous metastases [2,3]. Standard treatments for metastatic CRC (mCRC) are based on chemotherapy.

As is the case for many cancers, CRC staging is essential for optimal patient management. Accurate prognostication facilitates both therapeutic decisions and stratification in randomized clinical trials of cancer treatments. In CRC, the well-known TNM staging system is predominantly used [4]. In mCRC, two validated prognostic classification systems can be applied: Köhne prognostic index [5] for patients receiving front-line fluoropyrimidine monochemotherapy and GERCOR (Groupe Coopérateur Multidisciplinaire en Oncologie) prognostic index [6] for patients with oxaliplatin-based or irinotecan-based regimens. However, the models' ability to discriminate between patients on the basis of their prognosis (as measured by the C-index [7]) is still relatively modest. Thus, improvement of these prognostic indicators is required [6].

* Correspondence: diouf.momar@chu-amiens.fr
[1]Direction de la Recherche Clinique et de l'Innovation, Centre Hospitalier Universitaire Amiens, Nord, 1, Place Victor Pauchet, F-80054, Amiens, Cedex, France
[2]Methodology and quality of life in oncology Unit, EA 3181 CHU Besançon and the Qualité de Vie et Cancer clinical research platform, Dijon, France
Full list of author information is available at the end of the article

In palliative care patients, the prognostic value of health-related quality of life (QoL) has been demonstrated for several types of cancer [8-10]. For mCRC patients, QoL is known to be an independent prognostic factor for overall survival (OS) [8,11]. Hence, QoL is a candidate for the improvement of existing prognostic indices. Given

Table 1 Baseline demographic, clinical and laboratory variables for patients with and without available QoL data

		All patients		Available QoL		Missing QoL		All patients
Variable	Class	N	%	N	%	N	%	P
Age	≤65	353	57	138	55	215	58	
	>65	267	43	111	45	156	42	0.2900
Gender	Male	367	59	151	61	216	58	
	Female	252	41	98	39	154	42	0.5739
PS	0	333	54	122	49	211	57	
	1	239	38	110	44	129	35	
	2	48	8	17	7	31	8	0.0611
Number of sites	1	354	58	147	59	207	57	
	>1	260	42	102	41	158	43	0.5672
Liver involvement	No	149	24	52	21	97	27	
	Yes	460	76	197	79	263	73	0.0872
Metastases	Synchronous	415	68	168	68	247	68	
	Metachronous	196	32	80	32	116	32	0.9374
Adjuvant chemotherapy	No	488	79	200	81	288	78	
	Yes	130	21	48	19	82	22	0.4013
Tumour site	Colon	398	64	160	64	238	64	
	Rectum	211	34	86	35	125	35	
	both	11	2	3	1	8	1	0.6730
LDH	≤1xULN	380	61	134	56	246	66	
	>1xULN	240	39	115	44	125	34	0.0017
ALP	≤1xULN	350	56	129	52	221	60	
	>1xULN	270	44	120	48	150	40	0.0560
CEA	≤1xULN	177	28	61	25	116	31	
	>1xULN	443	72	188	75	255	69	0.0673
EuroQoL								
Mobility	1	223	81	223	81			
	2-3	54	19	54	19			
Self-care	1	255	93	255	93			
	2-3	19	7	19	7			
Usual activities	1	193	71	193	71			
	2-3	79	29	79	29			
Pain/discomfort	1	137	50	137	50			
	2-3	138	50	138	50			
Anxiety/depression	1	145	53	145	53			
	2-3	130	47	130	47			
VAS score				70 [10–100] **				

** Median (range).
ULN= Upper Limit of Normal.
VAS= visual analogue scale.
PS= performance status.
ALP= alkaline phosphatase.
LDH= serum lactate dehydrogenase.

that QoL is a multidimensional concept, there is a need to identify the QoL dimensions associated with OS for each specific type of cancer. The results of a recent study showed that social functioning (as measured with the EORTC QLQ-C30 tool) is an independent prognostic factor for survival in mCRC patients [12]. The objective of the present study was to assess the independent prognostic value of QoL in mCRC and evaluate its ability to improve the Köhne and GERCOR prognostic indices.

Methods

Patients

Individual patient data from the OPTIMOX1 phase III trial were analysed. The 620 evaluable patients from OPTIMOX1 were recruited from January 2000 to June 2002 from 56 institutions in five countries. In this trial, the oxaliplatin stop-and-go strategy proved to be as good as a continuous oxaliplatin-based chemotherapy strategy in previously untreated mCRC patients. The trial's inclusion and exclusion criteria are detailed elsewhere [13].

Quality of life assessment

Quality of life was self-reported by the patient using the generic EQ-5D questionnaire (also known as EuroQol) [14], which has five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) rated as one of three levels ("no problems",”some problems" and ”extreme problems", coded as 1, 2 and 3, respectively). The EQ-5D also includes a 100-centimetre visual analogue scale (VAS) for the self-assessment of overall health (0 = worst possible score; 100 = best possible score).

The GERCOR and Köhne prognostic indices

The Köhne prognostic index [5] comprises four variables: performance status (PS), number of metastatic sites, alkaline phosphatase (ALP) level and white blood cell (WBC) count. The GERCOR prognostic index [6] is based on two variables: PS and serum lactate dehydrogenase (LDH) level. Patients are classified into three risk groups (low, intermediate and high) in both models.

Statistical analysis

Demographic and clinical characteristics were summarized as frequency and percentage. In order to check whether selection bias was present, the patients’clinical characteristics were compared (with chi-squared test or Fisher's exact test) as a function of the available QoL data at baseline.

Overall survival was defined as the time from randomization to death (regardless of the cause) or last follow-up (censored data). All randomized patients with complete QoL data were included in the statistical analysis.

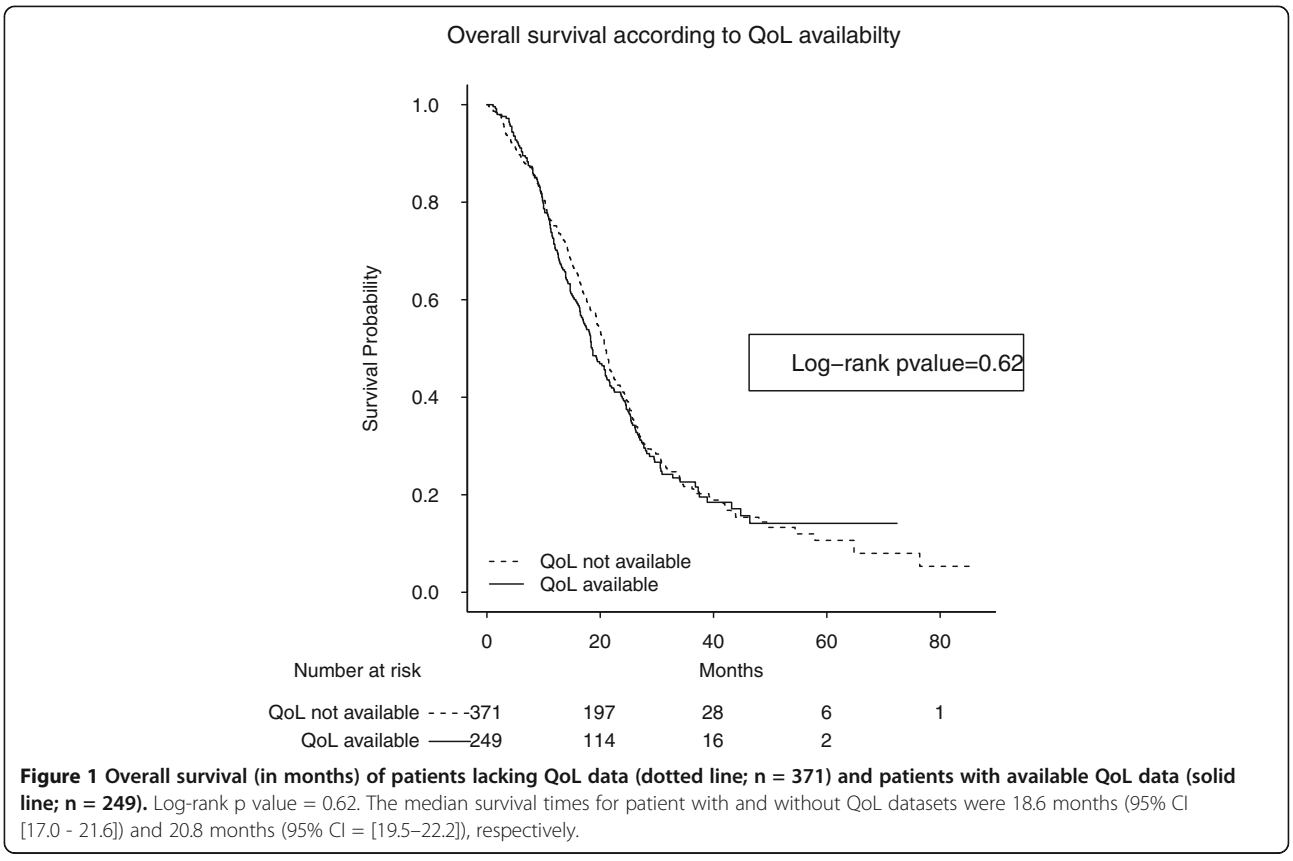


Figure 1 Overall survival (in months) of patients lacking QoL data (dotted line; n = 371) and patients with available QoL data (solid line; n = 249). Log-rank p value = 0.62. The median survival times for patient with and without QoL datasets were 18.6 months (95% CI [17.0 - 21.6]) and 20.8 months (95% CI = [19.5–22.2]), respectively.

Univariate and multivariate analysis were performed using Cox proportional hazards modelling, with calculation of the hazard ratio (HR) and the corresponding 95% two-sided confidence intervals (95%CI).

In order to evaluate the independent prognostic value of QoL, we built two multivariate models with backward selection. The first model included all demographic and clinical variables associated with OS ($p<0.1$) in univariate analysis. The second model included demographic, clinical and QoL variables with $p<0.1$ in univariate analysis.

Improvements in the prognostic index was evaluated by adding clinical variables (other than those used to

Table 2 Univariate and multivariate Cox analyses

		Univariate analysis			Multivariate analysis			Multivariate analysis		
					Model not including QoL			Full model, including QoL		
Variable	Class	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P
Age	≤65	1								
	>65	1.42	1.06 – 1.89	0.0178						
Gender	Male	1								
	Female	1.06	0.79 – 1.42	0.6945						
PS	0	1			1			1		
	1-2	1.84	1.38 – 2.46	<0.0001	1.98	1.44 – 2.73	<0.0001	1.87	1.35 – 2.59	0.0002
Number of sites	1	1			1			1		
	>1	1.47	1.10 – 1.97	0.0094	1.48	1.08 – 2.05	0.0160	1.48	1.07 – 2.04	0.0176
Liver involvement	No	1								
	Yes	1.14	0.795 – 1.65	0.4699						
Metastases	Synchronous	1								
	Metachronous	0.89	0.61 – 1.29	0.5403						
Adjuvant chemotherapy	No	1								
	Yes	0.95	0.76 – 1.19	0.68						
LDH	≤1xULN	1			1			1		
	>1xULN	2.04	1.48 – 2.80	<0.0001	1.93	1.39 – 2.68	<0.0001	1.83	1.31 – 2.55	0.0004
APL	≤1xULN	1								
	>1xULN	1.60	1.20 – 2.14	0.0016						
CEA	≤1xULN	1								
	>1xULN	1.48	1.01 – 2.18	0.0444						
EuroQoL										
Mobility	1	1						1		
	2-3	1.90	1.33 – 2.71	0.0004				1.66	1.12 – 2.48	0.0117
Self-care	1	1								
	2-3	1.52	0.88 – 2.62	0.1322						
Usual activities	1	1								
	2-3	1.20	0.88 – 1.64	0.2553						
Pain/discomfort	1	1								
	2-3	1.39	1.04 – 1.86	0.0239						
Anxiety/depression	1	1								
	2-3	1.45	1.09 – 1.93	0.0116						
VAS score		1.001	0.996 – 1.005	0.7975						
Harrell's C index						0.65 [0.61 – 0.69]			0.67 [0.63 – 0.71]	
						0.65*			0.66*	
Schemper statistic						9.32%			10.42%	

ULN = Upper Limit of Normal.
* = Optimism-corrected C-index.

build the prognostic index) and QoL variables (with p<0.1 in univariate analysis) to a model with backward selection (Köhne or GERCOR index being forced in the model). Patients with available QoL data for whom Köhne and GERCOR indices could be calculated were considered for prognostic systems' improvement.

The models were compared by calculating the Schemper statistic [15] and Harrell's C index [7]. The Schemper statistic is equivalent to R^2 in linear regression and quantifies the proportion of the survival variability that is explained by the model. Briefly, the higher the Schemper statistic is, the more accurate the OS predictions would be. Harrell's C index estimates discriminate capability, i.e. the ability to distinguish between high-risk and low-risk patients. The C-index varies from 0.5 (no discrimination) to 1 (perfect discrimination). Optimism-corrected C-index was calculated using 200 bootstrap replications.

Category-free net reclassification improvement [16] (NRI) was also calculated at various moments (12, 24 and 36 months), in order to evaluate the additional utility of QoL domains and other clinical factors. NRI quantifies"the correctness of upward and downward reclassification or movement of predicted probabilities as a result of adding a new marker". The confidence interval for NRI was calculated using the percentiles of 1000 bootstrap replications.

We also performed a sensitivity analysis using the multiple-imputation technique [17,18] (with 10 replications) for missing QoL data. The choice of 10 replications was prompted by the large amount of missing QoL data in the trial (60%). In line with Van Buuren's method [19], the demographic and clinical variables initially included in the final complete-data model, those associated with the lack of QoL data and those strongly associated with OS (albeit absent from the final model) were used as predictors for the imputation of missing QoL data using a logistic regression model (QoL coded as 2–3 vs. 1). Multiple imputation with 10 replications (of the original database) consisted in creating 10 plausible values for each missing data and thus generating 10 new complete databases. For each of the new databases, a standard analysis was performed and the results were combined into a single estimation of the parameter of interest, while taking account of the uncertainty of the imputation technique [20]. Variables selected more than 5 times out of 10 replications were included in the multivariate model after multiple imputations.

Since there was no within-imputation variance according to the Schemper statistic, the pooled estimate was presented as the median [range] [20].

Construction of the a modified prognostic index was based on linear transformation as follows: The regression

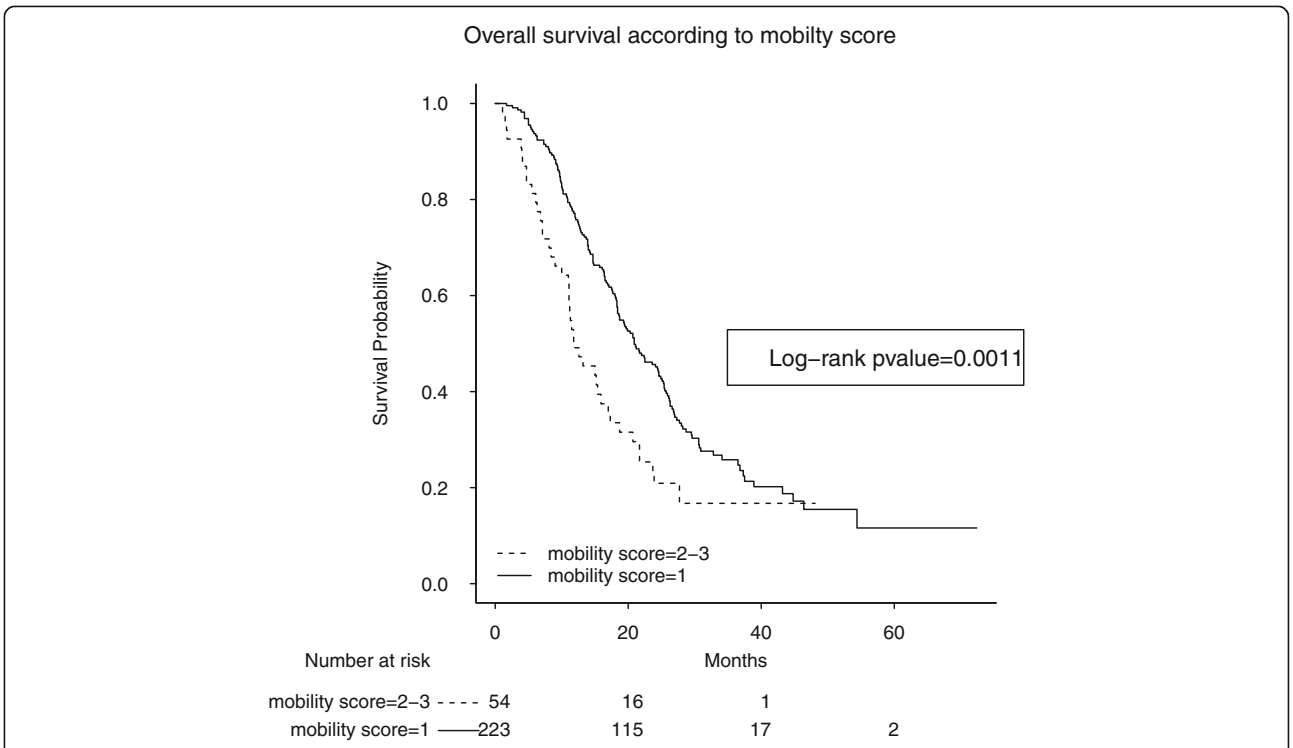


Figure 2 Overall survival (in months) of patients with mobility problems (as coded 2–3) (dotted line; n = 54) and patients without mobility problems (as coded 1) (solid line; n = 223). Log-rank p value = 0.0011. The median survival times were 20.9 (95% CI = [18.6–24.9]) months and 11.8 (95% CI = [11.1–17.3]) months for patients without problems (coded as 1) and those with problems (as coded 2–3), respectively.

Table 3 Improvement of Köhne prognostic index

Köhne prognostic index					
Variable	HR (95% CI)	P value	c-index	Schemper (%)	NRI (95% CI)
Köhne (2 vs. 1)	1.18 [0.96 – 1.47]	=0.1200			
Köhne (3 vs. 1)	2.66 [1.84 – 3.85]	<0.0001	0.54 [0.51 -0.57] *0.54	1.6	
Improvement of the Köhne prognostic index with clinical and QoL factors: complete-case analysis					
Köhne (2 vs. 1)	1.11 [0.80 – 1.55]	=0.5114			NRI at 12 months = 0.23 ([0.07; 0.46])
Köhne (3 vs. 1)	2.17 [1.25 – 3.75]	=0.0056			NRI at 24 months = 0.31 ([0.16; 0.44])
LDH (>1ULN vs. ≤ 1ULN)	2.09 [1.53 – 2.87]	<0.0001	0.67 [0.63 -0.71]	10.8	NRI at 36 months = 0.27 ([0.02; 0.50])
Mobility (2–3 vs. 1)	1.56 [1.05 – 2.32]	=0.0266	*0.66		
Pain/discomfort (2–3 vs. 1)	1.60 [1.17 – 2.18]	=0.0031			
Improvement of the Köhne prognostic index with clinical and QoL factors after multiple imputation					
Köhne (2 vs. 1)	1.24 [0.97 – 1.58]	=0.0780			
Köhne (3 vs. 1)	2.15 [1.43 – 3.24]	=0.0002			
LDH (>1ULN vs. ≤ 1ULN)	1.99 [1.61 – 2.46]	<0.0001	0.66 [0.59 -0.73]	8.63 [7.74 – 10.8]	
Mobility (2–3 vs. 1)	1.39 [1.06 – 1.83]	=0.0191	R = 65%		
Pain/discomfort (2–3 vs. 1)	1.67 [1.20 – 2.31]	=0.0031	R = 113%		

LDH = lactate dehydrogenase.
ULN = Upper Limit of Normal.
* = bootstrap C-index.
R = relative increase in variance due to missing data.
QoL = Quality of Life.
HR = Hazard ratio.
NRI = net reclassification improvement.
For multiple imputations, a logistic model was used: response variable = QoL scale (2–3 vs. 1) and exploratory variables were number of metastatic sites, liver involvement, WHO Performance Status, CEA, APL and LDH.
Variables considered in the imputation method (last model) were selected more than 5 times among the 10 replications of multiple imputations (see statistical method).

coefficient for each variable selected in the final multivariate complete case Cox model was divided by the lowest coefficient and rounded to the nearest integer [21]. The sum of these integers is the maximum score (M) for the modified index; hence the new score varied from zero to M. According to the score, the modified prognostic index was then arbitrary divided into three risk groups: good prognostic, intermediate prognostic and poor prognostic.

Survival distributions were estimated using the Kaplan-Meier method [22] and compared with the log-rank test.

All statistical analyses were carried out using SAS® software (version 9.2, SAS Institute Inc., Cary, NC) and R.2.12.0 software (free) using the Design, SurvIDINRI (for NRIs) and Multivariate Imputation by Chained Equations packages (http://www.multiple-imputation.com/). P-values were two-sided and variables with p<0.05 were considered significantly associated with OS in multivariate models.

Results

Patient characteristics

The patient baseline characteristics are summarized in Table 1, most of them were male (59%) and 43% were over the age of 65. Synchronous metastasis was predominant (68%) and most of the patients with metachronous metastasis received adjuvant chemotherapy (66%, 130/196).

Data on QoL was available for 249 of the 620 patients in the original OPTIMOX1 cohort (40%). Normal serum LDH was significantly more frequent in patients with missing QoL data. Patients with missing QoL data also tended to have lower serum ALP levels, a better PS and less liver involvement compared to patients with available QoL. Of the 249 patients, 75% died after a median follow-up period of 35.8 months (95% CI = [33.8–38.4]). There was no apparent correlation between the availability of QoL datasets and OS (Log-rank pvalue = 0.62; Figure 1).

Most of the patients had good QoL: 81%, 93%, 71%, 50% and 53% had no problems in terms of mobility, self-care, usual activities, pain/discomfort and anxiety/depression, respectively. The median VAS score was 70 (range = [10–100]).

Univariate analysis

Given that"extreme problems" (coded as 3) were infrequent, QoL item scores were pooled into two classes (i.e. a score of 1 vs. a score of 2 or 3). We also combined PS into 2 classes (0 vs. 1–2), due to the low proportion of patients with a PS score of 2.

Univariate analyses of clinical and QoL variables are summarized in Table 2. High serum LDH, poor PS, high serum ALP, >1 metastatic sites, age>65, high serum CEA, mobility problems (as coded 2–3) (Figure 2),

Table 4 Improvement of the GERCOR prognostic index

GERCOR prognostic index					
Variable	HR (95% CI)	P value	c-index	Schemper (%)	NRI (95% CI)
GERCOR (2 vs. 1)	1.82 [1.43 – 2.33]	<0.0001			
GERCOR (3 vs. 1)	3.10 [2.38 – 4.05]	<0.0001	0.63 [0.61 -0.66] *0.63	6.44	
Improvement of the GERCOR prognostic index clinical and QoL factors: complete-case analysis					
GERCOR (2 vs. 1)	1.70 [1.14 – 2.54]	=0.0090			NRI at 12 months = 0.35 [0.06; 0.44]
GERCOR (3 vs. 1)	3.35 [2.20 – 5.10]	<0.0001	0.67 [0.63 -0.71] *0.67	11.52	NRI at 24 months = 0.27 [0.04; 0.38]
					NRI at 36 months = 0.28 [0.01; 0.45]
Mobility (2–3 vs. 1)	1.77 [1.19 – 2.62]	=0.0047			
Anxiety/depression (2–3 vs. 1)	1.41 [1.03 – 1.92]	=0.0314			
Improvement of the GERCOR prognostic index clinical and QoL factors: multiple imputation					
GERCOR (2 vs. 1)	1.77 [1.36 – 2.30]	<0.0001			
GERCOR (3 vs. 1)	2.49 [1.84 – 3.38]	<0.0001			
ALP (>1ULN vs. ≤ 1ULN)	1.25 [1.00 – 1.57]	=0.0480	0.67 [0.64 -0.71]	9.56 [8.76 – 11.52]	
Mobility (2–3 vs. 1)	1.42 [1.08 – 1.86]	=0.0120	R = 60%		
Pain/discomfort (2–3 vs. 1)	1.55 [1.10 – 2.20]	=0.0140	R = 138%		

LD = lactate dehydrogenase.
ULN = Upper Limit of Normal.
* = bootstrap C-index.
R = relative increase in variance due to missing data.
QoL = Quality of Life.
HR = Hazard ratio.
NRI = net reclassification improvement.
For multiple imputations, a logistic model was used: response variable=QoL scale (2–3 vs. 1) and exploratory variables were number of metastatic sites, liver involvement, WHO Performance Status, CEA, APL and LDH.
Variables considered in the imputation method (last model) were selected more than 5 times among the 10 replications of multiple imputations (see statistical method).

pain/discomfort problems (as coded 2–3) and anxiety/depression problems (as coded 2–3) were associated with a poorer prognosis.

There were no significant associations between the risk of death and self-care (p = 0.1322), usual activities (p = 0.2553) and the VAS score (p = 0.1280) QoL scales on the other.

Multivariate analysis

The results for multivariate analyses are summarized on Table 2.

In the first model, high LDH, >1 metastatic sites and poor PS were associated with a shorter survival.

In the second model, high LDH, >1 metastatic sites, poor PS and mobility problems were associated with a shorter survival.

After multiple imputations, the pooled HR for mobility was 1.57 (95% CI = [1.16–2.12]) (p = 0.0043) in the model including LDH, the number of metastatic sites, PS, ALP, pain/discomfort and mobility (Additional file 1).

Improvement of prognostic indices

In order to evaluate improvements in performance of the Köhne and GERCOR prognostic indices, we first calculated their performance in our population (Table 3).

Improvement of the Köhne prognostic index

After addition of QoL and clinical variables to the Köhne prognostic index in a complete-case analysis (N = 236), high LDH, mobility and pain/discomfort problems appeared to be related to a shorter survival (Table 4). The

Table 5 Modified Köhne prognostic index

	0 point	1 point	2 points	3 points	4 points	5 points	6 points	7 points
Köhne	Köhne I	Köhne II						Köhne III
LDH	≤ 1ULN							>1ULN
Mobility score	1				2-3			
Pain/discomfort score	1				2-3			

The modified Köhne index varied from 0 to 22 points.
Poor prognosis: 15 to 22 points.
Intermediate prognosis: 8 to 14 points.
Good prognosis: 0 to 6 points.

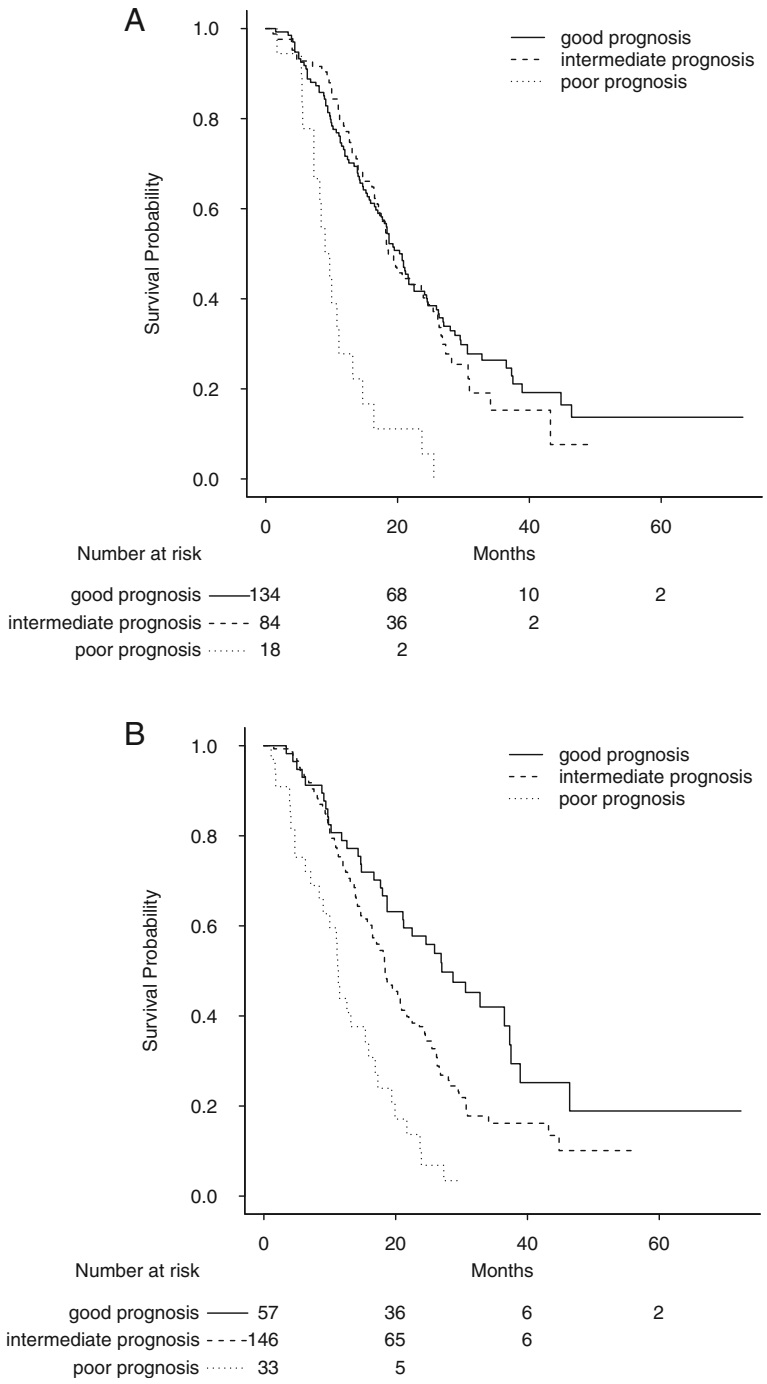


Figure 3 Survival strata according to the Köhne prognostic model before and after improvement. A: Overall survival (in months) for good, intermediate and poor prognosis according to the Köhne prognostic model. Median survival = 20.7 [17.7 – 24.4] for the group with good prognosis (n = 134); Median survival = 18.6 [17.1 – 25.4] for the group with intermediate prognosis (n = 84); Median survival = 9.0 [7.3 -14.7] for the group with poor prognosis (n = 18). Log-rank p = 0.0013. Optimism corrected C-index = 0.54. **B:** Overall survival (in months) for good, intermediate and poor prognosis according to the modified Köhne group. Median survival = 27.0 [21.1 – 37.5] for the group with good prognosis (n = 57); Median survival = 18.4 [16.5 – 21.6] for the group with intermediate prognosis (n = 146); Median survival = 11.3 [9.0 – 16.9] for the group with poor prognosis (n = 33). Log-rank p<0.0001. Optimism corrected C-index = 0.60.

C-index and Schemper statistic were improved while the NRIs were significantly different from zero (Table 3). A modified Köhne prognostic index was built using the above variables (Table 5).

Survival distributions for the Köhne and improved Köhne prognostic systems are shown in Figure 3 A&3B.

The Results of multiple imputations are summarized in Table 3.

A complete-case analysis of the GERCOR prognostic classification revealed that mobility and Anxiety/depression could improve performance: the C-index, Schemper statistic, and NRI are summarized in Table 4.

Based on these two new QoL scales, a modified GERCOR prognostic system was built using the above variables (Table 6).

Survival distributions for the GERCOR and improved GERCOR prognostic systems are shown in Figure 4A and Figure 4B.

The Results of multiple imputations are summarized in Table 4.

Discussion

In this study, EuroQol mobility dimension appeared to be the third most important prognostic factor (measured by the hazard ratio) for overall survival in unresectable mCRC, after serum LDH level and ECOG performance status. Self-reported QoL is known to be associated with OS in several types of cancer [8,9,11,12]. Our present results confirmed the independent prognostic value of QoL scales in patients with mCRC [8,11,12]. Our first multivariate model (including clinical and biochemical variables) revealed the prognostic value of LDH, PS and the number of metastatic sites, whereas our second model (with the addition of QoL) confirmed the value of LDH, PS and the number of metastatic sites and further identified the QoL"mobility" scale as an independent prognostic factor.

After multiple imputations, the mobility QoL scale remained significant despite its high associated relative increase in variance due to missing data imputation. Pain/discomfort was not significant but showed a prognostic value after the multiple- imputation analysis; this may be partially related to the high increase in variance due to missing QoL data.

We found that the Köhne prognostic system could be improved by including LDH, mobility and pain/discomfort in both complete-case and imputation analyses. Moreover, the GERCOR prognostic index was improved by mobility and anxiety/depression in a complete-case analysis and by ALP, mobility and pain/discomfort after multiple imputations. This difference in the selection of variables may be due to lack of power in the complete-case analysis albeit ALP was at the limit of statistical significance. Therefore the GERCOR prognostic index was essentially improved by QoL scales. The added value of QoL scales (completed by the patient) for improvements of the two prognostic systems revealed that the patient's perception of his/her disease was an important information to record for prognosis assessment in addition to the clinician's evaluation [23].

Despite a marked increase in variance due to missing data, the mobility and pain/discomfort QoL dimensions significantly improved the Köhne and GERCOR staging systems. This result comforted the independent prognostic value of these QoL scales in mCRC patients. The results for complete-case and multiple-imputation analysis were very similar. QoL significantly improved the prognostic indices with both methods (complete-case and multiple-imputation analyses). This may be related to the fact that the compete-case analysis was not biased. In fact, patients with and without QoL data at inclusion did not differ in terms of the median survival time [24] (i.e. missingness was not related to outcome).

It should be noted that such a large improvement in the C-index from 0.54 to 0.66 for the Köhne prognostic index has rarely been reported in prognostic studies. After the addition of both clinical and QoL factors, the NRIs were also statistically significant for both the Köhne and the GERCOR prognostic systems (95% CIs did not contained zero). The independent prognostic value of mobility and pain/discomfort QoL scales (using the EQ-5D) for mCRC is compatible with the result of Efficace [12] regarding the prognostic value of social functioning scale (using the EORTC QLQ-C30). In fact, mobility and pain problems could impair the social functioning QoL dimension.

One of the present study's strengths relates to its use of the easily understood and rapidly completed EQ-5D.

Table 6 Modified GERCOR prognostic index

	0 point	1 point	2 points	3 points	4 points
GERCOR	GERCOR I		GERCOR II	GERCOR III	
Mobility score	1	2-3			
Pain/discomfort score	1	2-3			

The modified GERCOR index varied from 0 to 5 points.
Poor prognosis: 4 to 5 points.
Intermediate prognosis: 2 or 3 points.
Good prognosis: 0 or 1 point.

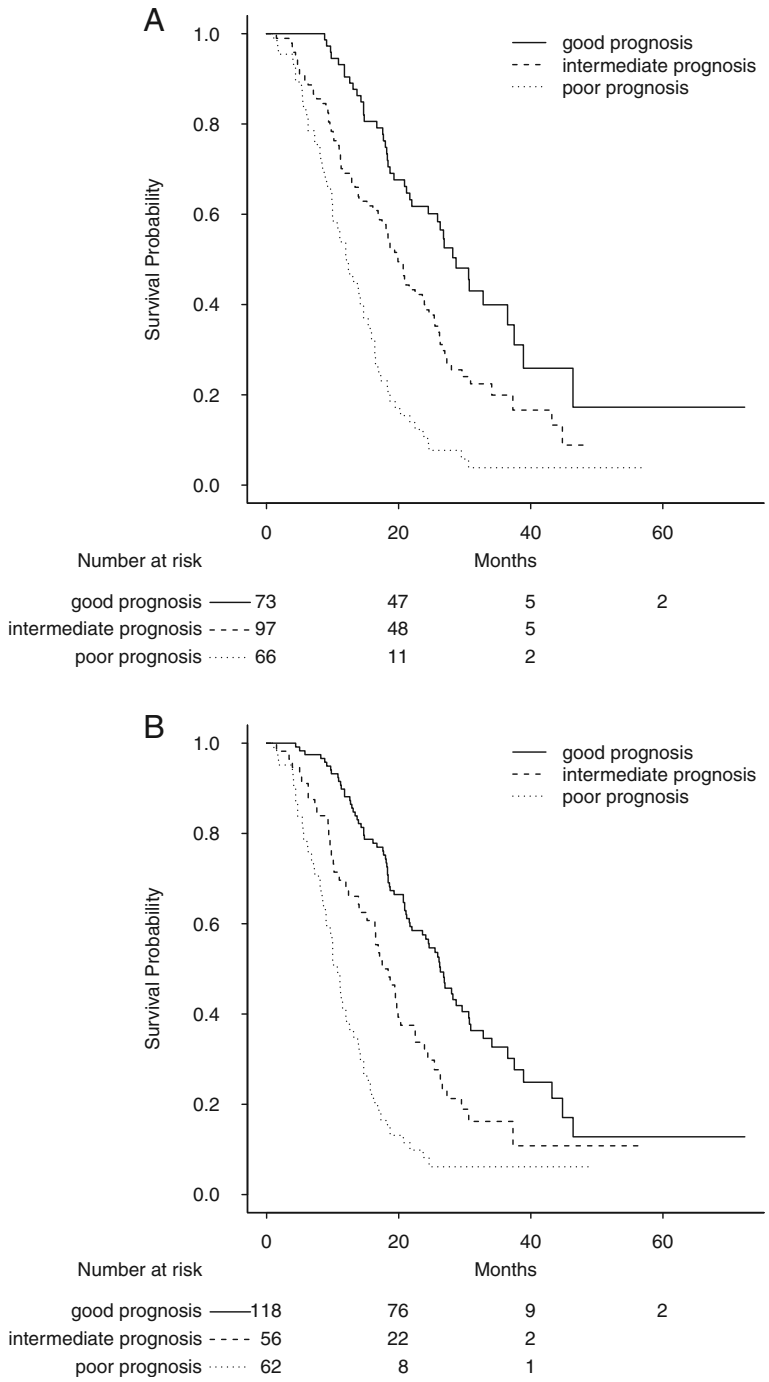


Figure 4 Survival strata according to the GERCOR prognostic model before and after improvement. A: Overall survival (in months) for good, intermediate and poor prognosis according to the GERCOR prognostic system. Median survival = 28.7 [24.5 – 38.9] for the group with good prognosis (n = 73); Median survival = 19.9 [18.1 – 23.9] for the group with intermediate prognosis (n = 97); Median survival = 12.1 [10.0 – 15.4] for the group with poor prognosis (n = 66). Log-rank p<0.0001. Optimism corrected C-index = 0.65. **B:** Overall survival (in months) for good, intermediate and poor prognosis according to the modified GERCOR prognostic system. Median survival = 28.2 [24.5 – 37.5] for the group with good prognosis (n = 68); Median survival = 21.6 [18.7 – 26.2] for the group with intermediate prognosis (n = 90); Median survival = 11.5 [10.0 – 14.7] for the group with poor prognosis (n = 78). Log-rank p<0.0001. Optimism corrected C-index = 0.66.

The EQ-5D was chosen because it was expected to be less time consuming and could prevent missing data. However, EQ-5D is not a cancer-specific questionnaire like the EORTC QLQ-C30 and it constitutes a limitation of our study. The high proportion of missing data (60%) and its large variability between countries (ranged from 5% to 66%) constitute another limitation in the generalizability of our results. Such a large heterogeneity in missing data might be related to the trial logistic and/or each country's culture. It is also important to note than our population came from a randomized controlled trial with restrictive inclusion and non inclusion criteria and might not be representative of mCRC patients in general [25]. Quality of Life may be an important parameter to record when assessing the situation of mCRC patients, since it improved the accuracy of OS prediction and greatly improved the two best-known prognostic classification systems for mCRC. We consider that QoL domains are important factors in the field of stratified therapy in the sense that knowing some aspect of the patient's self-reported QoL level could be decisive in the choice of different treatment options in the area of tailored medicine. By way of an example, a clinician might wish to avoid a treatment with pain as side-effect if the patient reported preexisting pain symptoms. Pain and mobility could also serve as an inclusion and/or stratification factor in randomized, controlled trials in mCRC.

Conclusion

Our results confirmed the prognostic value of QoL in mCRC patients. Thus, QoL scores should be recorded as it could give supplementary information to the clinician regarding the prognosis of a patient as well as in the judgment of an acceptable treatment side effect.

Additional file

Additional file 1: Results of the multivariate analysis after QoL imputation.

Abbreviations

QoL: Quality of life; mCRC: Metastatic colorectal cancer; OS: Overall survival; EQ-5D: Generic measure of health status developed by the EuroQol Group; HR: Hazard ratio; CI: Confidence interval; CRC: Colorectal cancer; TNM: Tumor Node Metastasis; GERCOR: Groupe Coopérateur Multidisciplinaire en Oncologie; EORTC: European Organization for Research and Treatment of Cancer; VAS: Visual analogue scale; PS: Performance status; ALP: Alkaline phosphatase; LDH: Lactate dehydrogenase; ITT: Intention to treat; NRI: Net reclassification improvement; CEA: Carcinoembryonic antigen; HDL: High density lipoprotein.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MD, BC, FB, AG, CL, CT, TF, NP, SD, AH, MH and MG the seven authors are justifiably credited with authorship, according to the authorship criteria. In detail: MD BC TF FB – conception, design, analysis and interpretation of data, drafting of the manuscript, final approval given; AG CL CT SD NP AH MH MG BC – acquisition of data, interpretation of data, critical revision of manuscript, final approval given. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Dr. David Fraser for advice in English language.

Author details

[1]Direction de la Recherche Clinique et de l'Innovation, Centre Hospitalier Universitaire Amiens, Nord, 1, Place Victor Pauchet, F-80054, Amiens, Cedex, France. [2]Methodology and quality of life in oncology Unit, EA 3181 CHU Besançon and the Qualité de Vie et Cancer clinical research platform, Dijon, France. [3]Department of medical oncology, Hôpital Saint-Antoine, Assistance Publique des Hôpitaux de Paris, UPMC Paris VI, Paris, France. [4]Oncology Multidisciplinary Research Group (GERCOR), Paris, France. [5]Biostatistics Unit, Claudius Régaud Institute, Toulouse, France. [6]Department of medical oncology, Hôpital Henri Mondor, Assistance Publique des Hôpitaux de Paris, Créteil, Paris, France. [7]Department of medical oncology, Institut Mutualiste Montsouris, Paris, France.

Received: 18 October 2013 Accepted: 29 April 2014
Published: 13 May 2014

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: **Global cancer statistics.** *CA Cancer J Clin* 2011, **61:**69–90.
2. Field K, Lipton L: **Metastatic colorectal cancer-past, progress and future.** *World J Gastroenterol* 2007, **13**(28):3806–3815.
3. Folprecht G, Grothey A, Alberts S, Raab HR, Köhne CH: **Neoadjuvant treatment of unresectable colorectal liver metastases: correlation between tumour response and resection rates.** *Ann Oncol* 2005, **16:**1311–1319.
4. Jessup JM, Gunderson LL, Greene FL, Washington MK, Compton CC, Sobin LH, Minsky B, Goldbert RM, Hamilton SR: **Staging System for Colon and Rectal Carcinoma.** *Ann Surg Oncol* 2010, **2011**(18):1513–1517.
5. Köhne CH, Cunningham D, Di Constanzo F, Glimelius B, Blijham G, Aranda E, Scheithauer W, Rougier P, Palmer M, Wils J, Baron B, Pignatti F, Schöffski P, Micheel S, Hecker H: **Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients.** *Ann Oncol* 2002, **13**(2):308–317.
6. Chibaudel B, Bonnetain F, Tournigand C, Bengrine-Lefevre L, Teixeira L, Artru P, Desramé J, Larsen AK, André T, Louvet C, de Gramont A: **Simplified prognostic model in patients with oxaliplatin-based or irinoecan-based first-line chemotherapy for metastatic colorectal cancer: A GERCOR study.** *Oncologist* 2011, **16**(9):1228–1238.
7. Harrell FE, Lee KL, Mark DB: **Tutorial in biostatistics. Multivariable prognostic models issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.** *Stat Med* 1996, **15:**361–387.
8. Montazeri A: **Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008.** *Health Qual Life Outcomes* 2009, **7:**102.
9. Quinten C, Coens C, Mauer M, Comte S, Sprangers MA, Cleeland C, Osoba D, Bjordal K, Bottomley A, EORTC Clinical Groups: **Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials.** *Lancet Oncol* 2009, **10:**865–871.
10. Diouf M, Filleron T, Barbare JC, Fin L, Picard C, Bouché O, Dahan L, Paoletti X, Bonnetain F: **The added value of quality of life (QoL) for prognoses of overall survival in patients with palliative hepatocellular carcinoma.** *J Hepatol* 2013, **58:**509–521.
11. Braun DP, Gupta D, Grutsch JF, Staren ED: **Can changes in health related quality of life scores predict survival in stage III and IV colorectal cancer?** *Qual Life Res* 2011, **9:**62.
12. Efficace F, Innominato PF, Bjarnason G, Coens C, Humblet Y, Tumolo S, Genet D, Tampellini M, Bottomley A, Garufi C, Focan C, Giacchetti S, Lévi F, Chronotherapy Group of the European Organisation for Research and Treatment of Cancer: **Validation of patient's self-reported social functioning as an independent prognostic factor for survival in metastatic colorectal cancer patients: results of an international study by the Chronotherapy Group of the European Organisation for Research and Treatment of Cancer.** *J Clin Oncol* 2008, **26:**2020–2026.

13. Tournigand C, Cervantes A, Figer A, Lledo G, Flesch M, Buyse M, Mineur L, Carola E, Etienne PL, Rivera F, Chirivella I, Perez-Staub N, Louvet C, André T, Tabah-Fisch I, de Gramont A: **OPTIMOX1: A randomized study of FOLFOX4 or FOLFOX7 with oxaliplatin in a stop-and-go fashion in advanced colorectal cancer—a GERCOR study.** *J Clin Oncol* 2006, **24:**394–400.
14. Group EQ: **EuroQol – a new facility for the measurement of health-related quality of life.** *Health Policy* 1990, **16:**199–208.
15. Schemper M: **Predictive accuracy and explained variation.** *Stat Med* 2003, **22:**2299–2308.
16. Pencina MJ, D'Agostino RB, Steyerberg EW: **Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers.** *Stat Med* 2011, **30:**11–21.
17. CLARK TG, Altman DG: **Developing a prognostic model in presence of missing data: an ovarian cancer study.** *J Clin Epidemiol* 2003, **56:**28–37.
18. Marshall M, Altman DG, Royston P, Holder RL: **Comparaison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study.** *BMC Med Res Methodol* 2010, **10:**7.
19. Van Buuren S, Boshuizen HC, Knook DL: **Multiple imputation of missing blood pressure covariates in survival analysis.** *Stat Med* 1999, **18:**681–694.
20. Marshall M, Altman DG, Holder RL, Royston P: **Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines.** *BMC Med Res Methodol* 2009, **9:**57.
21. Tournoux-Facon C, Paoletti X, Barbare JC, Bouché O, Rougier P, Dahan L, Lombard-Bohas C, Faroux R, Raoul JL, Bedenne L, Bonnetain F: **Development and validation of a new prognostic score of death for patients with hepatocellular carcinoma in palliative setting.** *J Hepatol* 2011, **54**(1):108–114.
22. Kaplan EL, Meier P: **Non parametric estimation from incomplete observations.** *J Am Stat Assoc* 1958, **53:**457–481.
23. Mauer M, Bottomley A, Coens C, Gotay C: **Prognostic factor analysis of health-related quality of life data in cancer: statistical methodological evaluation.** *Expert Rev Pharmacoecon Outcomes Res* 2008, **8**(2):179–196.
24. Vach W, Blettner M: **Missing data in epidemiologic studies**. In *Encyclopedia of Biostatistics.* New York: John Wiley & Sons; 1998:2641–2654.
25. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG: **Prognosis and prognostic research: what, why, and how?** *Br Med J* 2009, **338:**b375. 1317-1320.

doi:10.1186/1477-7525-12-69
Cite this article as: Diouf *et al.*: Could baseline health-related quality of life (QoL) predict overall survival in metastatic colorectal cancer? The results of the GERCOR OPTIMOX 1 study. *Health and Quality of Life Outcomes* 2014 **12**:69.

ii. Carcinome hépatocellulaire : valeur pronostique de la qualité de vie.

1) Résumé :

Rationnel: Plusieurs systèmes de classification pronostique ont été développés pour les patients atteints de CHC avancé. Une récente étude a montré que l'IP-OMS ajouté au système de classification CLIP donnait le meilleur indice de discrimination de Harrell. Nous avons étudié la valeur pronostique de la qdv pour les patients atteints de CHC et son intérêt pour améliorer la classification de ces patients.

Méthode: Nous avons analysé les données de l'essai CHOC avec évaluation du pouvoir discriminant sur la survie globale des systèmes de classification CLIP/GRETCH/BCLC/BoBar seuls et ensuite en association avec chacun des groupes de variables suivants : variables clinico-biologiques, scores de qdv en tant que variables continues, scores de qdv en tant que variables binaires, variables clinico-biologiques et scores de qdv en tant que variables continues, variables clinico-biologiques et scores de qdv en tant que variables binaires. La qdv juste avant la randomisation a été renseignée avec le questionnaire QLQ-C30 de l'EORTC. La performance des modèles a été évaluée avec l'indice de discrimination de Harrell et l'indice NRI.

Résultats: 271 patients ont été recrutés entre juillet 2002 et octobre 2003 dans 79 centres français. La qdv était renseignée pour 79% des patients (n=271). L'analyse univariée a montré que de meilleurs scores d'activité quotidienne (HR=0.991 [0.987–0.995]) et de fonction physique (0.991 [0.984–0.997]) étaient associés à une durée de survie plus longue. A l'inverse, de mauvais scores de fatigue (1.011 [1.006–1.015]) et de diarrhée (1.008 [1.002–1.013]) étaient associés à une durée de survie plus courte. Après ajustement par les paramètres démographiques et clinico-biologiques, seul un meilleur score d'activité quotidienne (0.993 [0.988–0.998]) était associé à une meilleure survie. L'addition des variables œdème, hépatomégalie, fatigue et diarrhée au système CLIP donnait la meilleure performance.

Conclusions: Les résultats confirment la valeur pronostique des scores de qdv pour la survie des patients atteints de CHC avancé. L'addition des scores de qdv améliore tous les systèmes de classification étudiés.

2) Article sur la qualité de vie dans le CHC:

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

The added value of quality of life (QoL) for prognosis of overall survival in patients with palliative hepatocellular carcinoma

Momar Diouf[1,2,*], Thomas Filleron[3], Jean-Claude Barbare[1,8,9], Loïc Fin[1], Carl Picard[1], Olivier Bouché[4], Laetitia Dahan[5], Xavier Paoletti[6,7], Franck Bonnetain[2]

[1]*Clinical Research and Innovation Directorate, Amiens University Hospital, Amiens, France;* [2]*Methodology and Quality of Life in Oncology Unit, EA 3181 CHU Besançon and the Qualité de Vie et Cancer Clinical Research Platform, Dijon, France;* [3]*Biostatistics Unit, Claudius Régaud Institute, Toulouse, France;* [4]*Department of Hepatogastroenterology, Reims University Hospital, Reims, France;* [5]*Department of Hepatogastroenterology and Digestive Oncology, La Timone University Hospital, AP-HM and the University of the Mediterranean, Marseille, France;* [6]*Biostatistics Service, Curie Institute, Paris, France;* [7]*INSERM U900, Paris, France;* [8]*Fédération Française de Cancérologie Digestive (French Society for Digestive Oncology), France;* [9]*Réseau des Investigateurs Pour le Carcinome Hépatocellulaire, France*

Background & Aims: Several prognostic classifications (PCs) have been developed for use in palliative care in patients with hepatocellular carcinoma (HCC). We have recently suggested that CLIP combined with WHO PS has the greatest discriminative power. We evaluated the prognostic value of quality of life (QoL) data and whether the latter could improve classification of palliative HCC patients.

Methods: This was a reanalysis from the CHOC trial with an evaluation of the discriminative power for overall survival (OS) of the established CLIP/GRETCH/BCLC/BoBar prognostic systems alone and then in association with each of the following groups of parameters: selected clinical factors, QoL as continuous variables, dichotomized QoL, selected clinical factors and continuous QoL, selected clinical factors and dichotomized QoL.

Baseline QoL was assessed using the EORTC QLQ-C30. Discriminative power was evaluated with the Harrell's C-index and net reclassification improvement.

Results: Quality of life was available in 79% of the patients (n = 271). Univariate analysis revealed that better role functioning (HR = 0.991 [0.987–0.995]) and better physical functioning (0.991 [0.984–0.997]) scores were associated with longer survival. In contrast, poorer score for fatigue (1.011 [1.006–1.015]) and diarrhoea (1.008 [1.002–1.013]) were associated with shorter survival. After adjustment for clinical and sociodemographic variables, only better role functioning score (0.993 [0.988–0.998]) was associated with longer survival. Adding oedema, hepatomegaly, fatigue and diarrhoea QoL scales to CLIP resulted in the best performance.

Conclusions: Our results confirm that QoL scales are independent prognostic factors of OS in palliative HCC patients. Incorporation of QoL data improved all the studied PCs.

Keywords: Quality of life; Prognosis; Hepatocellular carcinoma; Palliative; Improvement.

Received 6 August 2012; received in revised form 2 November 2012; accepted 12 November 2012; available online 22 November 2012

* Corresponding author. Address: Direction de la Recherche Clinique et de l'Innovation, Centre Hospitalier Universitaire Amiens Nord, 1, Place Victor Pauchet, F-80054 Amiens Cedex, France. Tel.: +33 322 668 139; fax: +33 322 667 911.

E-mail address: diouf.momar@chu-amiens.fr (M. Diouf).

Abbreviations: QoL, quality of life; PC, prognostic classification; HCC, hepatocellular carcinoma; EORTC, European Organization for Research and Treatment of Cancer; CLIP, Cancer of the Liver Italian Program; BCLC, Barcelona Clinic Liver Cancer; GRETCH, Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire; BoBar, BOnnetain & BARbare; OS, overall survival; WHO, World Health Organization; PS, performance status; AFP, alpha-fetoprotein; SD, standard deviation; ITT, intention-to-treat; 95% CI, 95% confidence interval; HR, hazard ratio; NRI, net reclassification improvement; HDL, high density lipoprotein; TKI, tyrosine-kinase inhibitor.

Introduction

Primary liver cancer is the fifth most common cancer and the third most common cause of cancer-related death in the world [1]. Hepatocellular carcinoma (HCC) is the main form of primary liver cancer [2] and about 70% of HCC patients are cared for in a palliative setting. In France, the main aetiology of HCC is alcohol abuse. Overall survival (OS) is poor, but can be improved by administration of one of the most recently developed treatments [3]. For patients in palliative care, the standard treatments are chemoembolization [4] and sorafenib [3]. Despite recent research results [5], the benefits of chemoembolization in HCC patients remain subject to debate. Hence, optimizing the treatment of HCC on the basis of the patient's characteristics is an important goal in a palliative setting and more generally.

One of the main objectives of a prognostic classification is to guide the selection of a therapeutic strategy according to the patient clinical, biochemical, and oncological characteristics. A classification can also be used to define eligibility criteria in randomized clinical trials and stratification criteria for randomization. Several prognostic classifications for HCC patients have been developed, including the Okuda staging system [6], the Cancer of the Liver Italian Program (CLIP) [7,8], the Barcelona Clinic Liver Cancer (BCLC) system [9] and the Groupe d'Étude et de Traitement du Carcinome Hépatocellulaire (GRETCH) system [10]. Several recent studies have emphasized the limitations of these scores in terms of discriminative power and OS prediction

Table 1. Definition of prognostic classifications.

Child-Pugh	Scores						
	0	1					
Presence of ascites	No	Yes					
Tumor size (>50%)	No	Yes					
Bilirubin (>50 μmol/L)	No	Yes					
Albumin (>30 g/L)	No	Yes					
CLIP	Scores						
	0	1	2				
Child-Pugh	A	B	C				
Tumor morphology	Uninodular and extension ≤50%	Multinodular and extension ≤50%	Massive or extension >50%				
AFP (>400 ng/d)	No	Yes					
Portal vein thrombosis	No	Yes					
BCLC	Scores						
	A1	A2	A3	A4	B	C	D
WHO PS	0	0	0	0	0	1-2	3-4
Tumor stage	Single	Single	Single	3 tumors <3 cm	Multinodular	Vascular invasion or extrahepatic spread	Any
Okuda	I	I	I	I-II	I-II	I-II	II
Liver functional status	No portal hypertension and normal bilirubin	Portal hypertension and normal bilirubin	Portal hypertension and abnormal bilirubin	Child-Pugh A-B	Child-Pugh A-B	Child-Pugh A-B	Child-Pugh C
GRETCH	Scores						
	0	1	2	3			
Bilirubin (≥50 μmol/L)	No			Yes			
Alkaline phosphatase (≥2N*)	No		Yes				
AFP (≥35 μg/L)	No		Yes				
Portal vein thrombosis	No	Yes					
Karnofsky (<80%)	No			Yes			
BoBar	Scores						
	0	1	2				
Non-small HCC	No	Yes					
Portal vein thrombosis	No	Yes					
Metastasis	No	Yes					
WHO PS	0	1	2-3				
Jaundice	No	Yes					
Ascites	No	Yes					
AFP (>200 μg/L)	No	Yes					
Alkaline phosphatase (>2N*)	No		Yes				

CLIP, stage I (0); stage II (1–2); stage III (3–5).
BCLC, stage A to D11.
GRETCH, A (0) B (1–5) C (6–11).
BoBar, Low risk (0–3); intermediate risk; (4–6) high risk (7–10).
WHO PS, World Health Organization's performance; AFP, alpha-fetoprotein.
*More than twice the upper limit of normal.

in a palliative setting (Colette *et al.* [11] and Tournoux-Facon *et al.* [12]). Hence, improving the quality and capabilities of these prognostic classifications remains an important challenge, since most patients have palliative HCC. To this end, Tournoux-Facon *et al.* suggested adding the World Health Organization's performance status (WHO PS) score to CLIP and further proposed a new prognostic classification (BoBar) that included metastasis, portal vein thrombosis, ascites status, tumour morphology, WHO PS, serum alpha-fetoprotein (AFP), jaundice and alkaline phosphatase [12]. These classifications were selected according to their discriminative ability (according to the C-index [13]) and the accuracy of the prognosis for the patient's individual outcome (according to the Schemper statistic [14]). As is the case for cancers in other sites (Quinten *et al.* [15]), we hypothesized that health-related quality of life (QoL) could improve the prediction of OS in palliative HCC.

Table 2. Patients' baseline characteristics. (A) Baseline demographic and clinical characteristics of patients with and without available QoL data. (B) Patients' baseline quality of life.

A	All patients (n = 271)		Patients with available QoL data (n = 215)		Patients lacking QoL data (n = 56)		*p* value
Variables	N	%	N	%	N	%	
Age (yr)							
≥ 65	180	66	145	67	35	62	
< 65	91	34	70	33	21	38	0.5901
Gender							
Male	202	75	165	77	37	66	
Female	69	25	50	23	19	34	0.1441
Cirrhosis							
Present	213	78	165	77	48	87	
Absent	55	20	48	22	7	13	
n.a.	3	2	2	1	1	2	0.1927
Portal vein thrombosis							
Yes	39	14	33	15	6	11	
No	232	86	182	85	50	89	0.3788
Extrahepatic metastasis							
Yes	63	23	47	22	16	27	
No	208	77	168	78	40	73	0.2896
Child-Pugh stage							
A	182	67	151	70	31	55	
B	64	24	51	24	13	23	
C	2	1	0	0	2	4	
D	23	8	13	6	10	18	0.0021
Non-small HCC							
Yes	89	33	70	33	19	34	
No	182	67	145	67	37	66	0.8458
Involved liver							
Volume ≤ 50%	224	83	179	83	45	80	
Volume > 50%	47	17	36	17	11	20	0.7549
Ascites							
Yes	44	16	32	15	12	21	
No	227	84	183	85	44	79	0.2368
Oedema							
Yes	40	15	27	13	13	23	
No	228	84	186	86	42	75	
n.a.	3	1	2	1	1	2	0.0750

(*continued on next page*)

Improving existing prognostic classifications by adding QoL could help physician optimize treatment for a given patient, in accordance with the goal of providing targeted, personalized therapy. A preliminary study by Bonnetain *et al.* [16] demonstrated the independent, prognostic value of self-reported QoL (assessed according to the Spitzer QoL Index) for HCC patients in a palliative setting and QoL's capability to improve HCC prognostic classifications, when compared with biochemical and/or clinical parameters.

In a population with a hepatitis B virus (HBV) aetiology, Yeo *et al.* [17] found that QoL scales rated with the EORTC QLQ-C30 were independent prognostic factors for OS in patients with unresectable HCC.

The objective of our present study was to confirm the prognostic value of QoL and to establish whether it could improve the performance of the CLIP, BCLC, GRETCH and BoBar classifications. In other words, the present study was designed to provide external validation of the results reported by Bonnetain *et al.* [16].

Patients and methods

Patients

Individual patient data were extracted from a phase III randomized, controlled trial (the CHOC trial) on the efficacy of long-acting octreotide in palliative HCC [18]. Between July 2002 and October 2003, 271 patients were randomized to receive either long-acting octreotide (n = 134) or placebo (n = 137). The CHOC trial failed to demonstrate the efficacy of octreotide in palliative HCC. The trial's inclusions criteria and results have been extensively described in detail elsewhere [18].

Quality of life assessment

Quality of life was assessed in the two weeks prior to randomization. Patients completed the European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 [19], a 30-item self-questionnaire with a 4-point Likert scale (''not at all''; ''a little''; ''quite a bit''; ''very much'') dealing with health and well-being in the previous two weeks. The 30 items are divided into 15 scales: global health, physical functioning, role functioning, emotional functioning, social functioning,

Table 2. *(continued)*

A	All patients (n = 271)		Patients with available QoL data (n = 215)		Patients lacking QoL data (n = 56)		*p* value
Variables	N	%	N	%	N	%	
Hepatomegaly							
Yes	162	60	132	61	30	53	
No	105	39	81	38	24	43	
n.a.	4	1	2	1	2	4	0.1849
Encephalopathy							
No	264	98	212	99	52	93	
Yes	2	1	0	0	2	3.5	
n.a.	5	2	3	1	2	3.5	0.0179
Jaundice							
Yes	19	7	12	6	7	12	
No	249	92	201	93	48	86	
n.a.	3	1	2	1	1	2	0.1126
WHO PS							
0	85	31	77	36	8	14	
1	123	45	99	46	24	43	
2	56	21	35	16	21	37	
n.a.	7	3	4	2	3	6	0.0002
Serum AFP (μg/L)							
<200	146	54	117	54	29	52	
≥200	125	46	98	46	27	48	0.7248
Serum albumin (g/L)							
<35	125	49	94	44	31	55	
≥35	132	51	114	53	18	32	
n.a	14		7	3	7	13	0.0021
Serum bilirubin (μmol/L)							
<20	143	53	122	57	21	37	
≥20	120	44	89	41	31	55	
n.a.	8	3	4	2	4	8	0.0082
Serum creatinine (μmol/L)							
≥80	139	51	119	55	20	36	
<80	126	47	91	42	35	62	
n.a.	6	2	5	3	1	2	0.0226
Serum alkaline phosphatase							
≤2N	201	74	166	77	35	62	
>2N*	62	23	47	22	15	27	
n.a.	8	3	2	1	6	11	0.0012

(*continued on next page*)

cognitive functioning, nausea, and vomiting, pain, fatigue, diarrhoea, insomnia, dyspnoea, appetite loss, and financial difficulties. The response for each scale was converted to a score ranging from 0 to 100 by using a linear transformation from the EORTC scoring manual [20]. For the last nine scales (i.e., the symptom scales), 100 is the worst score and 0 is the best score, whereas the opposite is true for the first six scales (i.e., those related to global health and functioning scales).

Prognostic classifications

The CLIP [8], BCLC [9], GRETCH [10] and BoBar [12] prognostic systems and scoring are described in Table 1. The Child-Pugh score [21] used in CLIP was based on ascites status, encephalopathy, total bilirubin level, prothrombin rate and albumin level.

Statistical analysis

For baseline demographic and clinical characteristics, categorical variables were summarized as the frequency and percentage, and continuous variables were summarized as the median and range or the mean ± SD. Clinical and medical patient characteristics were compared (in a Chi-square or Fisher's exact test) with the availability of a QoL questionnaire at baseline, in order to check whether selection bias was present.

Overall survival was defined as the time from randomization to death by any cause or the date of the last follow-up (censored data). Survival distributions were estimated using the Kaplan–Meier method and inter-group comparisons were performed with the log-rank test.

Median follow-up was estimated using the reverse Kaplan–Meier method.

All randomized patients with available QoL scores and clinical data were included in statistical analysis (i.e., as a modified intention-to-treat (ITT) population).

Since no pre-specified cut-off has been proposed for quality of life scales in HCC, the QoL dimensions were dichotomized into two levels ($\geqslant$50 *vs.* <50), in order to prevent overfitting.

Univariate and multivariate analyses were performed using Cox proportional hazard modelling to estimate hazard ratios. The corresponding 95% two-sided confidence intervals (95% CI) were also calculated. Proportional hazards assumptions were tested graphically.

Table 2. *(continued)*

A	All patients (n = 271)		Patients with available QoL data (n = 215)		Patients lacking QoL data (n = 56)		*p* value
Variables	N	%	N	%	N	%	
Okuda stage							
I	72	27	54	25	18	32	
II	187	69	153	71	34	61	
III	12	4	8	4	4	7	0.2394
CLIP score							
0	9	3	9	4	0	0	
1	26	10	19	9	7	13	
2	107	39	85	39	22	39	
3	94	35	74	34	20	36	
4	28	10	25	12	3	5	
5	7	3	3	2	4	7	0.0830
BCLC stage							
A	24	11	27	10	3	5	
B	32	15	39	14	7	13	
C	147	68	187	69	40	71	
D	12	6	18	6	6	11	0.3362
GRETCH							
A	49	23	57	21	8	14	
B	153	71	195	72	42	75	
C	13	6	19	7	6	11	0.2336

B			
QLQ-C30 scales	N	Median (min-max)	Mean ± standard deviation
Global health	234	67 (0-100)	61 ± 21
Physical functioning	237	80 (0-100)	74 ± 22
Emotional functioning	236	83 (0-100)	75 ± 29
Role functioning	233	83 (0-100)	72 ± 32
Cognitive functioning	238	83 (0-100)	83 ± 20
Social functioning	238	100 (0-100)	82 ± 25
Fatigue	237	33 (0-100)	39 ± 29
Nausea and vomiting	238	0 (0-100)	8 ± 19
Pain	237	17 (0-100)	26 ± 27
Dyspnoea	234	33 (0-100)	31 ± 32
Insomnia	232	33 (0-100)	31 ± 32
Appetite loss	234	0 (0-100)	24 ± 35
Constipation	233	0 (0-100)	17 ± 28
Diarrhoea	235	0 (0-100)	15 ± 25
Financial difficulties	234	0 (0-100)	6 ± 18

n.a., not available; WHO PS, World Health Organization's performance; AFP, alpha-fetoprotein.
*More than twice the upper limit of normal.

Using a complete-case analysis, we built a multivariable model (with backward selection) that included all clinical and QoL variables with a *p*-value below 10% in a univariate, available-case analysis.

Lastly, a multivariate, complete-case analysis was performed to identify any clinical variables and QoL scales that improved prognostic indicators. The following five-step algorithm (using Cox proportional hazards modelling with backward selection) was applied:

Step 1: evaluation of the performance (as defined below) of the four prognostic classifications.
Step 2: testing for improvement in the prognostic classifications by the inclusion of QoL as a continuous variable, by adding all QoL scales with a *p*-value <10% in univariate analysis.
Step 3: testing for improvement by the inclusion of dichotomized QoL, by adding all QoL scales with a *p*-value <10%.
Step 4: testing for improvement by the inclusion of clinical parameters not used to build the prognostic classifications and continuous QoL scales with a *p*-value <10%.
Step 5: testing for improvement by the addition of clinical parameters not used to build the prognostic classifications and dichotomized QoL scales with a *p*-value <10%.

Performance of the prognostic scores was assessed with the Schemper statistic [14] and Harrell's C-index [13]. The Schemper statistic is equivalent to R^2 in linear regression and quantifies the proportion of the survival variability that is explained by the model (i.e., the relative gain in predictive accuracy attributable to a given covariate). The higher the Schemper statistic, the more precise the individual prediction of overall survival. The Harrell's C-index estimates the proportion of correct predictions, i.e., the proportion of patients with better staging and

who had a better survival. The C-index varies from 0.5 (no discrimination) to 1 (perfect discrimination). Optimism-corrected C-statistics and the shrinkage factor [13] were calculated using 150 bootstrap replications.

Category-free net reclassification improvement (NRI) [22] was also calculated, in order to "quantify the correctness of upward and downward reclassification or movement of predicted probabilities as a result of adding a new marker". The NRI method combines measures of both discrimination and calibration. The 95% confidence interval for NRI was calculated using the percentiles of 1000 bootstrap replications.

We also performed a sensitivity analysis using a multiple imputation technique [23,24] (three replications) for missing QoL in an ITT analysis. Clinical variables associated with missing QoL were used as predictors in a logistic model for imputation. Multiple imputations with three replications consist in creating three plausible values for the missing data, which creates three complete databases. For each database, standard analysis is performed and the results are combined to yield a single estimation of the parameter of interest that takes into account the uncertainty of the imputation technique.

All analyses were carried out using SAS® version 9.2 (SAS Institute Inc., Cary, NC) and R.2.12.0 software packages (R.2.12.0 for the bootstrap and Schemper results). All reported *p*-values are two-sided. For multivariable models, variables with $p < 0.05$ were considered to be significantly associated with OS.

Results

Patients' characteristics

The patients' baseline characteristics are summarized in Table 2A. Most of the patients were male (75%) and were aged ⩾65 years (66%). Cirrhosis was present in 78% of the patients and 23% had extrahepatic metastasis. Eighty percent of the subjects have a good WHO PS (score 0–1). Most of the patients were Child-Pugh class A (67%), CLIP class 0–1–2 (52%), BCLC class C (68%) and GRETCH class B (71%).

Of the 271 patients randomized into the CHOC trial, 215 (79%) had a full set of baseline QoL data and formed the modified ITT population. High Child-Pugh stage, presence of encephalopathy, poor PS, high serum bilirubin, high serum alkaline phosphatase, and low serum creatinine were more frequent in patients with missing QoL data than in patients with available QoL data (Table 2A).

The QoL scales (range: 0–100 in all cases) are summarized in Table 2B and Fig. 1. The median global health scale score was 67 and the median social and physical functioning scores were 100 and 80, respectively. The worst symptom scales were fatigue, dyspnoea and insomnia, with a median of 33.

Overall survival

Univariate analysis

After a median follow-up period of 30.9 months (95% CI [30.0–34.4]), 249 patients had died (92%). The median OS time was 6.8 months (95% CI [5.8–7.9]). Patients without QOL data had significantly worse OS ($p < 0.0001$). The median OS times for patients with and without QoL data were 7.8 months (95% CI [6.8–9.7]) and 3.1 months (95%CI [2.1–6.4]), respectively (Fig. 2) .

Table 3 shows the results of univariate Cox regression analyses for clinical variables and each QLQ-C30 scale; clinical variables associated with OS were presence of cirrhosis, jaundice, hepatomegaly, oedema, ascites, metastasis, portal vein thrombosis, serum AFP ⩾200 μg/L, total bilirubin ⩾20 μmol/L, serum albumin <35 g/L, serum alkaline phosphatase above two upper limit of normal and poor PS were associated with shorter survival.

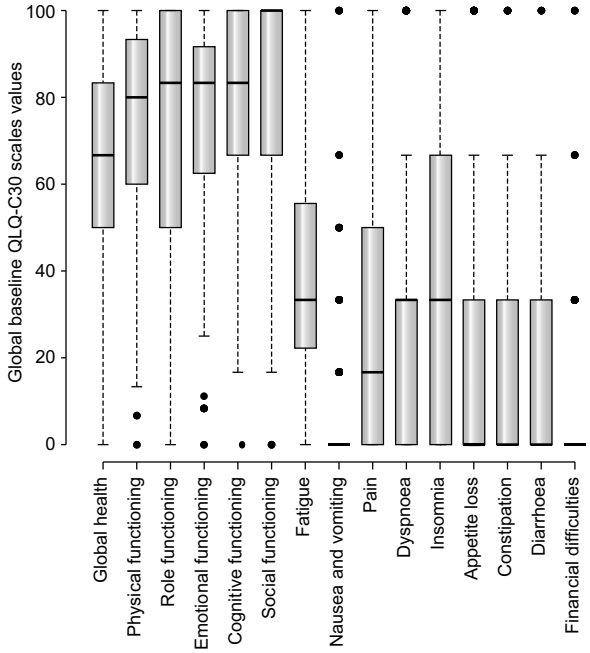


Fig. 1. Distribution of baseline QLQ-C30 scales values.

When QoL scales were analyzed as continuous factors, better score in global health, physical functioning, role functioning, and social functioning were associated with a longer survival. Poorer scores in fatigue, nausea, pain, dyspnoea, appetite loss and diarrhoea were associated with a worse survival. The optimism-corrected C-index ranged from 0.51 to 0.60 and Schemper statistic ranged from 0.34% to 3.61% for QoL scales.

After dichotomization of the QoL scales, patients with a poor score on the functioning scales (i.e., <50) were found to have a greater risk of death: this was the case for global health, physical functioning, role functioning and social functioning. A score >50 (reflecting a negative impact) on symptom scales was associated with worse survival, with significant relationships for fatigue, nausea, dyspnoea, appetite loss and diarrhoea. For dichotomized QLQ-C30 scales, the optimism-corrected C-index ranged from 0.52 to 0.57 and the Schemper statistic ranged from 0.6% to 2.11%.

The median OS times for patients with a good physical functioning score (⩾50) and those with a poor physical functioning score (<50) were 8.1 (95% CI = [6.9–9.9]) and 4.8 months (95% CI = [2.3–7.9]), respectively (Fig. 3). The median OS times for patients with a good fatigue score (<50) and those with a poor fatigue score (⩾50) were 8.9 (95% CI [7.3–10.3]) and 4.7 (95% CI [3.2–6.9]) months, respectively (Fig. 4).

Multivariate analysis

Table 3 shows the results of the multivariable Cox regression including clinical variables and continuous QoL scales identified in univariate analysis. In this model, a poorer role functioning score, serum AFP ⩾200 μg/L, total bilirubin >20 μmol/L, serum albumin <35 g/L, presence of portal vein thrombosis, distant metastasis, hepatomegaly, oedema and ascites were related to a shorter survival.

The C-index and Schemper statistic were 0.68 [0.64–0.72] and 14.61%, respectively.

When QoL was dichotomized, a poorer physical functioning score (<50) was associated with shorter survival.

Improvement of prognostic classifications

As shown in Table 4, the optimism-corrected C-index values for BCLC, GRETCH, BoBar, CLIP and CLIP+ PS were 0.57, 0.59, 0.63, 0.62 and 0.66, respectively. The corresponding Schemper statistic values were 2.97%, 5.05%, 8.83%, 6.32% and 8.79%, respectively (Table 4A–D).

Better scores (>50) in global health and the various functioning scales were associated with a longer survival whereas poorer scores in symptom scales (>50) were associated with a shorter survival.

All four prognostic classifications were improved by incorporation of the following QoL scales (treated as continuous variables), on the basis of a modified ITT analysis (Table 4A–D):

Table 3. Univariate and multivariate Cox analyses.

Variable	Univariate analysis				Multivariate analysis		
	Class	HR	95%CI	*p* value	HR	95%CI	*p* value
Age	>65	1.00					
	≤65	0.81	0.62-1.05	0.1109			
Sex	Female	1.00					
	Male	0.78	0.53-1.15	0.2158			
Cirrhosis	No	1.00					
	Yes	1.33	0.96-1.84	0.0847			
HCV	No	1.00					
	Yes	1.16	0.80-1.67	0.4310			
Jaundice	No	1.00					
	Yes	2.00	1.25-3.22	0.0040			
Hepatomegaly	No	1.00			1.00		
	Yes	1.27	0.98-1.64	0.0739	1.20	1.03-1.41	0.0217
Oedema	No	1.00			1.00		
	Yes	2.52	1.78-3.56	<0.0001	1.32	1.05-1.66	0.0168
Ascites	No	1.00					
	Yes	1.81	1.30-2.51	0.0004	1.62	1.08-2.43	0.0200
Tumor morphology	Uninodular and extension ≤50%	1.00					
	Multinodular and extension ≤50%	1.23	0.90-1.68				
	Diffuse or extension >50%	1.40	0.93-2.12	0.2462			
Metastasis	No	1.00			1.00		
	yes	1.46	1.09-1.96	0.0104	1.73	1.21-2.48	0.0026
Portal vein thrombosis	No	1.00			1.00		
	Yes	1.76	1.24-2.50	0.0015	2.61	1.72-3.96	<0.0001
AFP (µg/L)	<200	1.00			1.00		
	≥200	1.57	1.22-2.02	0.0004	1.70	1.25-2.32	0.0008
Total bilirubin (µmol/L)	<20	1.00			1.00		
	≥20	1.92	1.49-2.47	<0.0001	1.55	1.13-2.13	0.0069
Albumin (g/L)	≥35	1.00			1.00		
	<35	1.56	1.20-2.00	0.0006	1.44	1.06-1.94	0.0185
Serum creatinine (µmol/L)	≥80	1.00					
	<80	1.06	0.83-1.37	0.6340			
Alkaline phosphatases	≤2N	1.00	1.36-2.45	<0.0001			
	>2N**	1.82					
WHO PS	0	1.00					
	1	1.43	1.06-1.91				
	2-3	2.00	1.40-2.85	0.0006			
Global health	*Continuous	0.991	0.985-0.996	0.0013			
	≥50	1.00					
	<50	1.61	1.16-2.27	0.0044			
Physical functioning	*Continuous	0.991	0.984-0.997	0.0035			
	≥50	1.00			1.00		
	<50	1.82	1.28-2.63	0.0010	2.00	1.32-3.04	0.0012
Role functioning	*Continuous	0.991	0.987-0.995	<0.0001			
	≥50	1.00					
	<50	1.47	1.05-2.08	0.0263			

(*continued on next page*)

Table 3. (*continued*)

Variable		Univariate analysis			Multivariate analysis		
	Class	HR	95%CI	*p* value	HR	95%CI	*p* value
Emotional functioning	*Continuous	0.992	0.987-0.998	0.0058			
	≥50	1.00					
	<50	1.39	0.95-2.04	0.0892			
Cognitive functioning	*Continuous	1.00	0.994-1.007	0.9665			
	≥50	1.00					
	<50	0.97	0.48-1.96	0.9322			
Social functioning	*Continuous	0.99	0.985-0.996	0.0005			
	≥50	1.00					
	<50	1.85	1.19-2.86	0.0069			
Fatigue	*Continuous	1.011	1.006-1.015	<0.0001			
	<50	1.00					
	≥50	1.60	1.20-2.15	0.0015			
Nausea and vomiting	*Continuous	1.008	1.001-1.014	0.0225			
	<50	1.00					
	≥50	1.73	1.02-2.93	0.0419			
Pain	*Continuous	1.008	1.003-1.013	0.0007			
	<50	1.00					
	≥50	1.29	0.95-1.75	0.1047			
Dyspnoea	*Continuous	1.006	1.002-1.010	0.0059			
	<50	1.00					
	≥50	1.36	1.00-1.86	0.0482			
Insomnia	*Continuous	1.002	0.998-1.006	0.2279			
	<50	1.00					
	≥50	1.18	0.87-1.60	0.2747			
Appetite loss	*Continuous	1.006	1.002-1.009	0.0038			
	<50	1.00					
	≥50	1.54	1.11-2.13	0.0100			
Constipation	*Continuous	1.003	0.998-1.009	0.1767			
	<50	1.00					
	≥50	1.28	0.87-1.89	0.2095			
Diarrhoea	*Continuous	1.008	1.002-1.013	0.0060			
	<50	1.00					
	≥50	1.83	1.18-2.83	0.0070			
Financial difficulties	*Continuous	1.006	0.999-1.013	0.0942			
	<50	1.00					
	≥50	1.48	0.73-3.01	0.2764			

*QoL (Quality of life) scales were analyzed as continuous variables.
**More than twice the upper limit of normal.
Cox multivariable model with QoL as a continuous variable: C-index multivariate = 0.68 (CI [0.64–0.72]).

CLIP, BCLC and GRETCH could be improved with fatigue and diarrhoea QoL scales. BoBar was improved by addition of dyspnoea and diarrhoea.

After exploration of the added value of clinical factors and continuous QoL for improvement of the prognostic scores, CLIP plus oedema, hepatomegaly, fatigue and diarrhoea remained the best prognostic score. It had a corrected C-index of 0.68, an NRI of 0.74 ([0.56; 1.19)] and a Schemper statistic of 13.39% (Table 4A).

The improvements with dichotomized QOL scales were as follows (Table 4A–D):

- CLIP plus physical functioning and global health
- BCLC plus fatigue and diarrhoea
- GRETCH plus physical functioning and diarrhoea
- BoBar plus dyspnoea and diarrhoea.

After exploration of the added value of clinical factors and dichotomized QoL for improvement of the prognostic scores, CLIP plus oedema, physical functioning, hepatomegaly, global health and diarrhoea remained the best prognostic score. It had the best discriminant ability (C-index = 0.67) and the greatest Schemper statistic (13.09%). The NRI for our improved CLIP (0.55 [0.31; 0.79]) was as high as those found for GRETCH and BCLC.

These results remained valid after 150 bootstrap operations (Table 4A–D). Furthermore, the corrected slope shrinkage was 0.89, 0.95, 0.92 and 0.98 for BCLC, GRETCH, CLIP and BoBar, respectively, thus indicating good calibration and little or no overfitting.

In sensitivity analysis (multiple imputations) including all 271 patients, CLIP was improved by oedema, physical functioning and global health in the first replication, whereas BoBar was improved by dyspnoea and diarrhoea (Supplementary Table 1).

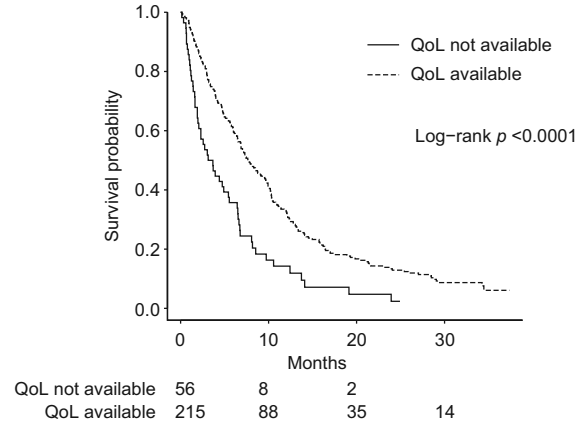


Fig. 2. Overall survival (months) of patients with lacking QoL (solid line; n = 56) and patients with available QoL (dotted line; n = 215). Log-rank p <0.0001.

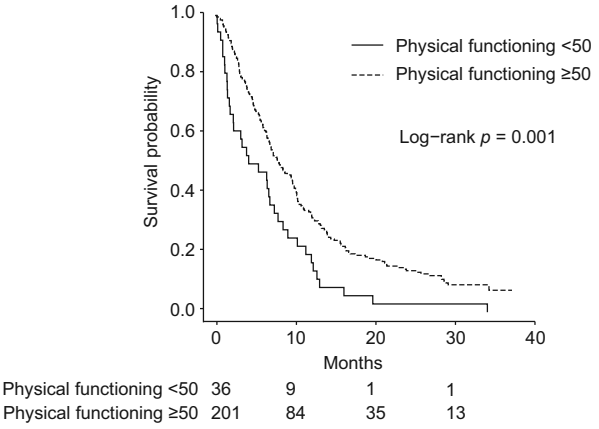


Fig. 3. Overall survival (months) of patients with poor physical functioning score (solid line; n = 31) and patients with good physical functioning score (dotted line; n = 201). Log-rank p = 0.001.

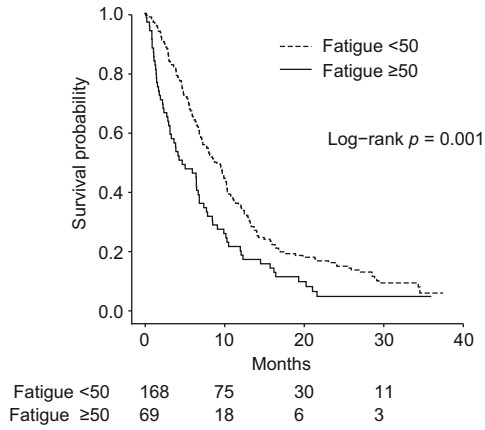


Fig. 4. Overall survival (months) of patients with poor fatigue score (solid line; n = 69) and patients with good fatigue score (dotted line; n = 168). Log-rank p = 0.0015.

Discussion

We established that role functioning, fatigue and diarrhoea QoL scales (as assessed by the EORTC QLQ-C30) were independent prognostic factors of OS in patients with palliative HCC. In multivariable Cox analysis, in which QoL scores were treated as continuous variables, role functioning was the main independent prognostic factor (in addition to clinical variables). Moreover, addition of QoL scales improved all prognostic classifications: fatigue and diarrhoea were frequently selected when QoL scales were analysed as continuous variables. Although classification of HCC patients is crucial for clinicians, there is no consensus on which staging system should be used for HCC patients in the palliative setting. Given that the statistical performance levels of existing prognostic indicators are modest [11,12], it is essential to improve these prognostic scores and provide patients with the most appropriate, targeted treatment. Our present results confirmed the prognostic and predictive value of QoL scales in palliative HCC patients as a means of improving prognostic classifications in this setting [15,16].

We found that (i) CLIP was improved by oedema, hepatomegaly, fatigue and diarrhoea, (ii) BCLC was improved by oedema, portal vein thrombosis, serum AFP, alkaline-phosphatase, fatigue and diarrhoea, (iii) GRETCH was improved by oedema, fatigue and diarrhoea, and (iv) BoBar was improved by dyspnoea and diarrhoea. Regardless of the prognostic classifications studied, fatigue and diarrhoea were the continuous QoL dimensions that yielded the greatest improvement.

To the best of our knowledge, CLIP (alone or supplemented with the WHO PS) currently appears to be the best classification for patients with palliative HCC [12]. Our results highlighted a modest improvement in the discriminant ability of our optimized CLIP (improved by oedema, hepatomegaly, fatigue and diarrhoea), when compared with CLIP+ WHO PS (0.68 *vs.* 0.66, respectively). However, the explained variation was clearly greater (13.39% *vs.* 8.79%). The apparently modest improvement in the C-index when comparing CLIP + clinical factors + QoL scales and CLIP+ WHO PS (from 0.66 to 0.68) may nevertheless be clinically relevant. Indeed, Pepe [25] showed that a small improvement in the C-index may be associated with the addition of a clinical factor with a hazard ratio of 3 or more. Moreover, for the prediction of cardiovascular disease, Cook [26] showed that after accounting for known factors (age, smoking status and blood pressure), the addition of serum HDL resulted in an improvement of 1% in the C-index (0.76–0.77). Pencina [22] also suggested that (i) the C-index is overly two conservative for the quantification of usefulness of addition of a new biomarker to a predictive model and (ii) NRI is a more intuitive measure of model performance.

Most of the confidence intervals for our NRIs did not include zero; this highlighted the performance improvements for performance of the prognostic classifications other than BoBar (for which the improvement was not trivial since the lower limit of confidence interval was highly negative).

When QoL scales and potential clinical variables were added to CLIP, the WHO PS was no longer significant, and fatigue and diarrhoea became significant. This result is in agreement with the findings of Osoba [27] and with Quinten *et al.* meta-analysis [15], showing that QoL scales and WHO PS were highly correlated and that self-reported fatigue and diarrhoea QoL scales were more informative than clinician-reported WHO PS. Mauer *et al.*

Table 4. Statistical performance of the four prognostic classifications. (A) Improvement of CLIP, (B) improvement of BCLC, (C) improvement of GRETCH, (D) improvement of BoBar.

A	**CLIP**				
		HR (95% CI)	Schemper (%)	C-index (95% CI) and *NRI (95% CI)	Bootstrap C-index
CLIP without improvement		(1.64 [0.76-3.52], II *vs.* I) (3.47 [1.61-7.47], III *vs.* I])	6.32	0.63 [0.58; 0.65]	0.62
CLIP + PS	CLIP PS	(1.60 [0.74-3.43], II *vs.* I) (3.33 [1.55-7.17], III *vs.* I]) (1.36 [1.01-1.82], 1 *vs.* 0) (1.86 [1.30-2.65], 2 *vs.* 0])	8.79	0.66 [0.62; 0.69] *	0.66
CLIP + continuous QoL scales	CLIP Fatigue Diarrhoea	(1.32 [0.61-2.86] , II *vs.* I) (2.89 [1.33-6.27] , III *vs.* I]) (1.010 [1.005-1.015]) (1.006 [1.00-1.012])	10.57	0.68 ([0.63; 0.72])	0.67
CLIP + dichotomized QoL scales	CLIP GH PF	(1.55 [0.71-3.39], II *vs.* I) (3.94 [1.79-8.63], III *vs.* I]) (1.71 [1.18-2.46]) (2.01 [1.35-3.01])	10.54	0.66 ([0.62; 0.70])	0.65
CLIP + clinical + QoL (continuous scales)	CLIP Oedema Hepatomegaly Fatigue Diarrhoea	(1.01 [0.46-2.23], II *vs.* I) (2.45 [1.12-5.34], III *vs.* I]) (2.02 [1.34-3.04]) (1.49 ([1.01-2.04]) (1.010 [1.005-1.015]) (1.007 [1.001-1.012])	13.39	0.68 ([0.64; 0.72]) *0.74 ([0.56; 1.19]) **0.40 ([-0.06; 0.76])	0.68
CLIP + clinical + QoL (dichotomized scales)	CLIP Oedema Hepatomegaly PF GH Diarrhoea	(1.17 [0.52-2.60], II *vs.* I) (3.20 [1.45-7.07], III *vs.* I]) (1.78 [1.18-2.71]) (1.43 ([1.05-1.94]) (1.92 [1.27-2.90]) (1.59 [1.09-2.33]) (1.63 [1.00-2.64])	13.09	0.68 ([0.64; 0.72]) *0.55 ([0.31; 0.79])	0.67

B	**BCLC**				
		HR (95% CI)	Schemper (%)	C-index (95% CI) and *NRI (95% CI)	Bootstrap C-index
BCLC without improvement		(1.08 [0.63-1.83], B *vs.* A) (1.72 [1.11-2.66], C *vs.* A) (3.11 [1.66-5.81], D *vs.* A)	2.97	0.58 [0.54; 0.61]	0.57
BCLC + continuous QoL scales	BCLC Fatigue Diarrhoea	(1.01 [0.57-1.78], B *vs.* A) (1.44 [0.91-2.29], C *vs.* A) (2.71 [1.34-5.47], D *vs.* A) (1.009 [1.004-1.014]) (1.006 [1.00-1.012])	5.96	0.62 ([0.58; 0.66])	0.62
BCLC + dichotomized QoL scales	BCLC Fatigue Diarrhoea	(1.00 [0.58-1.76], B *vs.* A) (1.51 [0.95-2.41], C *vs.* A) (2.76 [1.36-5.58], D *vs.* A) (1.44 [1.06-1.96]) (1.75 [1.10-2.75])	4.50	0.60 ([0.56; 0.64])	0.59
BCLC + clinical + QoL (continuous scales)	BCLC Oedema PVT AFP Alk-phos Fatigue Diarrhoea	(1.14 [0.64-2.03], B *vs.* A) (1.60 [1.00-2.57], C *vs.* A) (2.45 [1.17-5.14], D *vs.* A) (2.47 [1.58-3.87]) (2.06 [1.38-3.08]) (1.88 [1.40-2.52]) (1.56 [1.11-2.21]) (1.008 [1.003-1.013]) (1.006 [1.00-1.011])	13.84	0.68 ([0.65; 0.72]) *0.60 ([0.28; 0.83])	0.67
BCLC + clinical + QoL (dichotomized scales)	BCLC Oedema PVT AFP Alk-phos PF Diarrhoea	(1.13 [0.63-2.03], B *vs.* A) (1.69 [1.06-2.71], C *vs.* A) (2.62 [1.26-5.44], D *vs.* A) (2.41 [1.55-3.76]) (1.97 [1.32-2.95]) (2.03 [1.50-2.73]) (1.65 [1.16-2.33]) (1.74 [1.17-2.58]) (1.74 [1.08-2.81])	12.62	0.67 ([0.63; 0.71]) *0.56 ([0.32; 0.88])	0.66

(*continued on next page*)

Table 4. *(continued)*

C		**GRETCH**			
		HR (95% CI)	Schemper (%)	C-index (95% CI) and *NRI (95% CI)	Bootstrap C-index
GRETCH without improvement		(1.76 [1.28-2.42], B *vs.* A) (5.07 [2.94-8.75], C *vs.* A)	5.05	0.59 [0.56; 0.62	0.59
GRETCH + continuous QoL scales	GRETCH Fatique Diarrhoea	(1.59 [1.13-2.24], B *vs.* A) (4.58 [2.44-8.60], C *vs.* A) (1.008 [1.003-1.013]) (1.007 [1.001-1.012])	7.81	0.64 ([0.60; 0.68])	0.63
GRETCH + dichotomized QoL scales	GRETCH PF Diarrhoea	(1.81 [1.28-2.55], B *vs.* A) (5.21 [2.77-9.80], C *vs.* A) (1.60 [1.10-2.35]) (1.94 [1.21-3.12])	6.21	0.61 ([0.57; 0.64])	0.60
GRETCH + clinical + QoL (continuous scales)	GRETCH Oedema Fatigue Diarrhoea	(1.56 [1.10-2.21], B *vs.* A) (4.41 [2.32-8.37], C *vs.* A) (2.18 [1.45-3.28]) (1.008 [1.003-1.013]) (1.007 [1.001-1.013])	9.95	0.66 ([0.62; 0.70]) *0.35 ([-0.01; 0.76])	0.65
GRETCH + clinical + QoL (dichotomized scales)	GRETCH Oedema PF Diarrhoea	(1.76 [1.24-2.50], B *vs.* A) (5.09 [2.70-9.61], C *vs.* A) (2.20 [1.47-3.29]) (1.59 [1.08-2.34]) (1.95 [1.21-3.14])	8.53	0.63 ([0.59; 0.67]) *0.56 ([0.41; 0.68])	0.63
D		**BoBar**			
		HR (95% CI)	Schemper (%)	C-index (95% CI) and *NRI (95% CI)	Bootstrap C-index
BoBar without improvement		(2.33 [1.77-3.06], 2 *vs.* 1) (4.94 [3.03-8.05], 3 *vs.* 1])	8.83	0.63 [0.60; 0.67]	0.63
BoBar + continuous QoL scales	BoBar Dyspnoea Diarrhoea	(2.37 [1.76-3.21], 2 *vs.* 1) (4.21 [2.40-7.40], 3 *vs.* 1]) (1.005 [1.001-1.010]) (1.007 [1.001-1.012])	10.21	0.67 [0.63; 0.71]	0.66
BoBar + dichotomized QoL scales	BoBar Dyspnoea Diarrhoea	(2.39 [1.77-3.23], 2 *vs.* 1) (4.35 [2.50-7.57], 3 *vs.* 1]) (1.67 [1.05-2.67]) (1.50 [1.08-2.08])	9.67	0.66 ([0.62; 0.70])	0.65
BoBar + clinical + QoL (continuous scales)	BoBar Dyspnoea Diarrhoea	(2.35 [1.73-3.19], 2 *vs.* 1) (4.24 [2.40-7.47], 3 *vs.* 1) (1.005 ([1.001-1.010]) (1.007 [1.001-1.012])	10.21	0.67 ([0.63; 0.71]) *0.26 ([-0.18; 0.60])	0.66
BoBar + clinical + QoL (dichotomized scales)	BoBar Diarrhoea	(2.12 [1.56-2.87], 2 *vs.* 1) (4.13 [2.37-7.19], 3 *vs.* 1) (1.85 [1.16-2.93])	8.89	0.63 ([0.60; 0.67]) *0.18 ([-0.88; 0.35])	0.65

QoL, health-related quality of life; HR, hazard ratio; PC, prognostic classification; PS, performance status; PF, physical functioning; GH, global health; PVT, portal vein thrombosis; AFP, alpha-fetoprotein; Alk-phos, alkaline phosphatase.
*NRI, Net reclassification improvement (compared to the original PC without additional variables); **NRI, Net reclassification improvement (compared to CLIP + PS).

[29] hypothesized that self-reported fatigue and diarrhoea QoL scales might account for the patient's early perception of the severity of disease and would thus extend the time interval for predicting OS. Thus, the patient is the person best able to evaluate his/her health and well-being and can provide important information for predicting OS.

Most of the variables included in the final model were QoL domains, which appeared to be more informative than clinical parameters, although other clinical variables may be useful for improving prognostic classifications. Our results confirmed that QoL can now be considered as a valuable, relevant parameter for improving prognostic classifications. Oedema (present in the multivariable analysis) was added to three prognostic scores and showed its independent prognostic value for HCC patients. One of the strengths of this study is that the QLQ-C30 is a well-recognized QoL assessment tool; its multidimensional aspect enables different levels of the patient's well-being and health to be evaluated.

When QoL was analyzed as a dichotomized variable for improving prognostic scores, we found that CLIP was improved by oedema, hepatomegaly, physical functioning, global health and diarrhoea; BCLC was improved by oedema, portal vein thrombosis, alkaline-phosphatase, physical functioning and diarrhoea; GRETCH was improved by oedema, physical functioning and diarrhoea and, lastly, BoBar was improved by diarrhoea. The observed disparities between dichotomized and continuous analyses of the selected QoL scales confirm the influence of empirical dichotomization in our results. Our empirical choice of a cut-off of 50 for the different QoL scales (which reduces the amount of information derived from QoL) constituted a study

limitation. From a methodological point of view, it would have been better to use QoL as a continuous variable; this would have improved the prognostic classification and yielded a more accurate prediction of OS [28,29]. Our results confirmed this point (Table 4). However, in clinical practice, it would be more difficult to calculate a score derived from a continuous variable. In general, prognostic classification including continuous biological parameters has used dedicated cut-offs derived from clinical practice (as in prognostic classifications for HCC [6,8–10] and metastatic colorectal cancer [30,31] and the lymphoma prognostic index [32]). To promote clinical uptake, we adopted a conservative approach in which the *a priori* empirical cut-off of 50 (rather than the median or another value derived from the QoL score distribution) prevents overfitting.

Again, role functioning was the most informative QoL scale when analyzed as a continuous variable, but did not make a significant contribution when dichotomized. Despite the variable selection procedure (89% of the patients with a good role functioning score had good physical functioning score; $p = 3.88.10^{-12}$ in a Chi-squared test), this disparity might be due to the empirical choice of the 50-point cut-off for all QoL scales. In fact, the cut point might be more appropriate for physical functioning than for role functioning. The bootstrap C-index was very similar to the C-index found in multivariable analysis, suggesting that our results have internal validity.

Further research will seek to determine and validate the optimal cut-off for the different QoL scales by applying an appropriate methodology [33,34].

Complete baseline QoL datasets were available in 79% of the patients. Some of the clinical variables differed when compared between patients with and without QoL data, suggesting the presence of selection bias and limiting the prognostic ability of QoL. Patients lacking QoL data had poorer survival than the others. This finding suggests that patients who are healthy enough to complete the QoL assessment may not be representative of the target population. To take this problem into account, we performed a sensitivity analysis by imputing missing QoL data (multiple imputations [23,24] with three replications). The improvements in prognostic scores by QoL in the sensitivity analysis were very similar to those in the complete-case analysis. Hence, the differences in the selected variables might be related to a lack of power in the complete-case analysis. This would inflate standard errors and affect the significance of variables in model selection [35]. Given that the WHO PS was the main explanatory variable for multiple imputations and was low for patients lacking QoL data, the imputed QoL value for these patients would also be low. As expected, the prognostic classifications performed better after imputation.

Our results confirmed that QoL dimensions should be taken into account when building new prognostic classifications for palliative HCC. Our results also demonstrated that QoL measurements before treatment need to be used in both clinical practice and clinical research in this patient population. In fact, QoL may enable physicians to better classify patients and identify the best individual therapeutic options. However, it should be noted that our population came from a randomized clinical trial with strict inclusion criteria, which limited our ability to generalize the results to all patients with palliative HCC [36]. Given the substantial number of tests performed, there is a need to externally validate the improvements prognostic classifications' performance in a similar population by using the EORTC-QLQ-C30. Addition of the EORTC QLQ-HCC18 disease-specific questionnaire [37] may optimize selection of the most prognostic QoL scales, since scales derived from this tool are specific to our population and may capture prognostic information more accurately.

Our study highlighted the need to integrate QoL into palliative HCC clinical trials, either by including QoL in the eligibility criteria or using a QoL score as a stratification factor for randomized studies. Quality of life could also be integrated into clinical decision-making. By way of example, for a Child B/C patient with major deterioration in predictive QoL scales, TKI-based treatment may be debatable. Indeed, diarrhoea and fatigue were often reported as adverse events of treatment [3]. Poorer scores in QoL scales related to these two symptoms are associated with a shorter survival. If these symptoms are already present, aggravation by treatment should be closely monitored. We suggest that treatment probably should come along supportive care with regular control of QoL.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhep.2012.11.019.

References

[1] Parkin DM, Bray F, Ferlay J, et al. Estimating the world cancer burden: Globocan 2000. Int J Cancer 2001;94:153–156.
[2] Kew MC. Hepatocellular cancer. A century of progress. Clin Liver Dis 2000;4:257–268.
[3] Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2008;359:378–390.
[4] Llovet JM, Real MI, Montaña X, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. Lancet 2002;359:1734–1739.
[5] Doffoël M, Bonnetain F, Bouché O, et al. Multicentre randomised phase III trial comparing tamoxifen alone or with transarterial lipiodol chemoembolisation for unresectable hepatocellular carcinoma in cirrhotic patients. Eur J Cancer 2008;44:528–538.
[6] Okuda K, Ohtsuki T, Obata H, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer 1985;56:918–928.
[7] The Cancer of Liver Italian Program Investigators. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: The Cancer of the Liver Italian Program (CLIP) investigators. J Hepatol 1998;28:751–755.
[8] The Cancer of Liver Italian Program Investigators. Prospective validation of the CLIP score: a new prognostic system for patients with cirrhosis and hepatocellular carcinoma. J Hepatol (Baltimore, Md.) 2000;31:840–845.
[9] Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–338.
[10] Chevret S, Trinchet JC, Mathieu D, et al. A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d'Etude et de Traitement du Carcinome Hepatocellulaire. J Hepatol 1999;31:133–141.
[11] Collette S, Bonnetain F, Paoletti X, et al. Prognosis of advanced hepatocellular carcinoma: comparison of three staging systems in two French clinical trials. Ann Oncol 2008;19:1117–1126.

[12] Tournoux-Facon C, Paoletti X, Barbare JC, et al. Development and validation of a new prognostic score of death for patients with hepatocellular carcinoma in palliative setting. J Hepatol 2011;54:108–114.
[13] Harrell FE, Lee KL, Mark DB. Tutorial in biostatistics. Multivariable prognostic models issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statist Med 1996;15:361–387.
[14] Schemper M. Predictive accuracy and explained variation. Statist Med 2003;22:2299–2308.
[15] Quinten C, Coens C, Mauer M, et al. Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. Lancet Oncol 2009;10:865–871.
[16] Yeo W, Mo FKF, Koh J, et al. Quality of life is predictive of survival in patients with unresectable hepatocellular carcinoma. Ann Oncol 2006;17:1083–1089.
[17] Bonnetain F, Paoletti X, Collette S, et al. Quality of life as a prognostic factor of overall survival in patients with advanced hepatocellular carcinoma: results from two French clinical trials. Qual Life Res 2008;17:831–843.
[18] Barbare JC, Bouche O, Bonnetain F, et al. Treatment of advanced hepatocellular carcinoma with long-acting octreotide: a phase III multicentre, randomised, double blind placebo-controlled study. Eur J Cancer 2009;45:1788–1797.
[19] Aaronson NK, Ahmedzai S, Bergman B, 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 1993;Inst. 85:365–376.
[20] Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. On behalf of the EORTC quality of life group: the EORTC QLQ-C30 scoring manual. third ed. Brussels, Belgium: European Organization for Research and Treatment of Cancer; 2001.
[21] Pugh RNH, Murray-Lyon IM, Dawson JL, et al. Transection of the oesophagus for bleeding oesophageal varices. Br J Surg 1973;60:646–664.
[22] Pencina Michael J, D'Agostino Ralph B, Steyerberg Ewout W. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statist Med 2011;2011:11–21.
[23] Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Statist Med 1999;18:681–694.
[24] Clark TG, Altman DG. Developing a prognostic model in presence of missing data: an ovarian cancer study. J Clin Epidemio 2003;56:28–37.
[25] Pepe M, Janes H, Longton G, et al. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol 2004;159:882–890.
[26] Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928–935.
[27] Osoba D. What has been learned from measuring health-related quality of life in clinical oncology? Eur J Cancer 1999;35:1565–1570.
[28] Mauer M, Bottomley A, Coens C, et al. Prognostic factor analysis of health-related quality of life data in cancer: statistical methodological evaluation. Expert Rev Pharmacoecon Outcomes Res 2008.
[29] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127–141.
[30] Köhne CH, Cunningham D, Di CF, et al. Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. Ann Oncol 2002;13:308–317.
[31] Chibaudel B, Bonnetain F, Tournigand C, et al. Simplified prognostic model in patients with oxaliplatin-based or irinoecan-based first-line chemotherapy for metastatic colorectal cancer: a GERCOR study. The Oncologist 2011;16:1228–1238.
[32] Solal-Céligny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. Am Soc Hematol 2004;104:1258–1265.
[33] Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and application to decision-making about cancer treatments. Statist Med 2000;19:113–132.
[34] Faraggi D, Simon R. A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis. Statist Med 1996;15:2203–2213.
[35] Marshall M, Altman DG, Royston P, et al. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. BMC Med Res Methodol 2010;10:7.
[36] Moons KGM, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? Br Med J 2009;338:1317–1320, b375.
[37] Blazeby JM, Currie E, Zee BC, et al. Development of a questionnaire module to supplement the EORTC QLQ-C30 to assess quality of life in patients with hepatocellular carcinoma, the EORTC QLQ-HCC18. Eur J Cancer 2004;40:2439–2444.

iii. Carcinome hépatocellulaire : valeurs seuil pour la qualité de vie.

1) Résumé :

Rationnel: La valeur pronostique de la qdv est maintenant validée pour les patients atteints de cancer à un stade avancé. Cependant, pour être utilisés en routine clinique les scores de qdv doivent être dichotomisés. Pour les scores de qdv, les valeurs seuil sont souvent arbitraires et basées sur les percentiles. Le but de cette étude est d'identifier des valeurs seuil optimales pour six scores de qdv et de quantifier le gain de performance associé à l'ajout de la qdv pour quatre scores de classification publiés.

Méthode: Les données analysées proviennent de 271 patients recrutés entre juillet 2002 et octobre 2003 dans 79 centres en France dans le cadre de l'essai CHOC sur l'efficacité de l'octréotide-retard pour le CHC avancé. Les données de qdv étaient recueillies à l'aide du questionnaire QLQ-C30 de l'EORTC avec des scores allant de 0 à 100.

La détermination des valeurs seuil optimales a été réalisée par la méthode de Faraggi.

L'amélioration de la performance des systèmes de classification pronostique a été évaluée à l'aide l'indice de discrimination de Harrell et les indices NRI et IDI.

Résultats: Les valeurs seuil optimales étaient de 50, 58.33, 66.66, 66.66, 0 et 33.33 pour le score de santé globale, de fonction physique, d'activité quotidienne, de fatigue, de dyspnée et de diarrhée, respectivement.

L'addition des scores de qdv et des paramètres clinco-biologiques a amélioré tous les systèmes de classification étudiés. L'étendue de l'augmentation de l'indice de Harrell, de l'indice NRI à 3 mois, de l'indice IDI à 3 mois était de [0.02; 0.09], [0.24; 0.78] et [0.02; 0.10], respectivement.

Conclusion: Ces valeurs seuil pour les scores de qdv pourraient être utiles pour identifier un sous-groupe de patient de très mauvais pronostic, améliorant ainsi la planification des essais cliniques et l'adaptation des traitements.

2) Article sur les valeurs seuil de qualité de vie dans le CHC:

The Oncologist
The official journal of the Society for Translational Oncology

September 25, 2014

Momar Diouf, M.D.
Direction de la Recherche Clinique et de l'Innovation
Centre Hospitalier Universitaire Amiens Nord
1, Place Victor Pauchet
F-80054 Amiens cedex, France

E-mail: diouf.momar@chu-amiens.fr

RE: MS# T14-175PR R1

Dear Dr. Diouf:

We are pleased to inform you that your revised manuscript, "Optimal Cut-Points for QLQ-C30 Scales: Utility for Clinical Trials and Updates of Prognostic Systems in Advanced Hepatocellular Carcinoma," has been reviewed by our Section Editors and has been accepted for publication. Page proofs will be sent to you in the near future. Please return your corrected proofs within 48 hours to avoid any delays in publication.

We are pleased to provide ***The Oncologist*** as the vehicle for this paper's international publication.

Yours sincerely,

Bruce A. Chabner, MD
Editor-in-Chief

NOTES FROM EDITORIAL OFFICE

The Oncologist accompanies each review article and research article with an impact statement called "Implications for Practice." This is a summary statement that describes the overarching significance of the article. This brief statement (approximately 100 words) should clearly define the potential impact and importance of the manuscript, especially in terms of its clinical applications. Whenever possible, it should be written in language comprehensible by a lay reader. Please email this statement to the Editorial Office: EditorialOffice@TheOncologist.com.

AlphaMed Press • 318 Blackwell St. Suite 260 • Durham, North Carolina 27701 USA
Telephone: 919-680-0011 • Fax: 919-680-4411 • TheOncologist@AlphaMedPress.com
AlphaMed Europe, Ltd. • Queen's University-Belfast • Centre for Cancer Research and Cell Biology • 97 Lisburn Road • Belfast, Northern Ireland, UK BT9 7BL

Optimal cut-points for QLQ-C30 scales: Utility for clinical trials and updates of prognostic systems in advanced hepatocellular carcinoma.

Sort title: Quality of life cutoff in Hepatocellular Carcinoma.

Momar Diouf[1, 2], Franck Bonnetain[2], Jean-Claude Barbare[1, 6], Olivier Bouché[3], Laetitia Dahan[4], Xavier Paoletti[5], Thomas Filleron[7].

[1] Clinical Research and Innovation Directorate, Amiens University Hospital, Amiens, France.

[2] Methodology and quality of life in oncology Unit, EA 3181 CHU Besançon and the *Qualité de Vie et Cancer* clinical research platform, Dijon, France.

[3] Department of Hepatogastroenterology, Reims University Hospital, Reims, France.

[4] Department of Hepatogastroenterology and Digestive Oncology, La Timone University Hospital, AP-HM and the University of the Mediterranean, Marseille, France.

[5] Biostatistics Service, Curie Institute, Paris, France

[6] Fédération Française de Cancérologie Digestive (French Society for Digestive Oncology).

[7] Biostatistics Unit, Claudius Régaud Institute, Toulouse, France.

Corresponding author:
Momar Diouf
Direction de la Recherche Clinique et de l'Innovation. Centre Hospitalier Universitaire Amiens Nord, 1, Place Victor Pauchet, F-80054 Amiens cedex, France.
E-mail: **diouf.momar@chu-amiens.fr**

Tel: +33 3 22 668 139
Fax: +33 3 22 667 911

Words count including references: 3113

Total number of table: 4

Total number of figures: 3

Supplementary figures: 2

Abstract

Background:

Health-related quality of life (QoL) is now validated as a prognostic factor for cancer patients. However, to be used in routine practice, QoL scores must be dichotomized. Cutoff points are usually based on arbitrary percentile values. We aimed to identify optimal cutoff points for six QoL scales and quantify their added utility in the performance of four prognostic classifications in hepatocellular carcinoma patients.

Methods:

We reanalyzed data of 271 patients with advanced HCC recruited between July 2002 and October 2003 from 79 institutions in France in the CHOC trial designed to assess the efficacy of long-acting octreotide. QoL was assessed with the EORTC QLQ-C30 questionnaire and the scores ranged from 0 to 100.

Identification of optimal cutoff points was based on the Faraggi method.

Improvement in the performance of prognostic classifications was studied by Harrell's C-index, net reclassification (NRI) and integrated discrimination improvement (IDI).

Results:

We found that 50, 58.33, 66.66, 66.66, 0 and 33.33 were optimal cutoff points for global health, physical functioning, role functioning, fatigue, dyspnea and diarrhea, respectively.

The addition of QoL and clinical factors improved the performance of all four prognostic classifications: the ranges of the improvement in C-index, 3-month NRI and IDI were [0.02; 0.09], [0.24; 0.78] and [0.02; 0.10], respectively.

Conclusion:

These cutoff values for QoL scales can be useful to identify HCC patients with very poor prognosis, thus improving the design of clinical trials and treatment adjustment for these patients.

Keywords: Optimal cutoff point - Quality of life – Hepatocellular carcinoma –- Prognostic classification.

This abstract was presented at the ESMO annual meeting 2012 in Vienna: number 1430P.

Introduction:

Primary liver cancer is the sixth most common cancer and the third most fatal cancer in the world [1]. Hepatocellular carcinoma (HCC) accounts for 90% of all primary liver cancers and only about 30% [2] of newly diagnosed HCC patients are eligible for curative treatment (liver transplantation, liver resection or radiofrequency ablation). Patients with intermediate HCC receive transcatheter arterial chemoembolization (TACE) while the standard of care for patients with advanced HCC is Sorafenib.

Prognostic value of health-related quality of life (QoL) has already been validated for palliative HCC patients [3, 4, 5] and for other cancer types [6]. Quality of life is often assessed by means of a questionnaire comprising several items. The patient's response to all of these items is converted to domain-specific scores that can be considered to be quantitative continuous variables.

It is well known that the use of continuous variables is statistically preferable to the use of categorized variables in prognostic studies [7]. However, to be easily used in routine staging, QoL measures must be categorized into a smaller number of levels. Like other laboratory parameters included in HCC prognostic indices (albumin, bilirubin, alpha-fetoprotein etc.), physicians usually based their decisions on a binary normal/abnormal assessment: to treat vs. not to treat. Median, percentiles or other arbitrary values have been selected as cutoffs for dichotomization into good or poor prognosis in the majority of studies [8]. Other commonly used methods are visual inspection of scatter plots [9, 10] and systematic search for the cut-point associated with a minimum chi-squared *p*-value [9, 10]. These less rigorous methods of categorization resulted in a marked heterogeneity of cut-points in the medical literature. An example of this possible heterogeneity in cutoff points was illustrated in Altman's prognostic study [12] in breast cancer in which he found 19 different cut-points for S-phase (phase of the cell cycle in which Deoxyribonucleic acid replication occurs).

Dichotomization of QoL scales could also facilitate their use when defining eligibility criteria or stratification factors for studies of new treatments. Another way of using QoL scales would be to add them to existing prognostic systems for HCC patients: Cancer of the Liver Italian Program (CLIP) [13], Barcelona Clinic Liver Cancer (BCLC) [14], Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire (GRETCH) [15] and BOnnetain & BARbare prognostic index (BoBar) [16]. The first

three staging systems were originally developed for all HCC patients and their limits for prognostic assessment in advanced HCC patients (corresponding to the BCLC class-C) have been underlined [17, 18]. Consequently, no consensus has been reached concerning the best prognostic system to be used for advanced HCC [18]. Adding QoL scales to existing prognostic indices could improve the physician's management of the patient's disease and could also allow the patient's perception of health to be taken into account in order to achieve stratified therapy. As the above prognostic systems were built with categorized variables, QoL scales need to be dichotomized before being included in these prognostic classifications. This simplified interpretation of QoL can only be achieved at a price: loss of information, as values close to the cutoff point but in opposite directions are treated as equally different as the minimum and maximum value of the continuous variable. Furthermore, a cutoff of point equal to the median value (which is not necessarily the optimal cutoff point) is equivalent to losing one third of the data, thereby resulting in loss of statistical power [7]. To limit this loss of power, we propose:

a) To determine the optimal cut-points (if they exist) using the method described by Faraggi et al. [19] for the six most statistically significant European Organization for Research and Treatment of Cancer QLQ-C30 scales (when these scales were treated as continuous variables) for overall survival prediction in a population of patients with palliative HCC: global health, physical functioning, role functioning, fatigue, dyspnea and diarrhea [3]. The existence of an optimal cut-point must be interpreted as follows: A point that divides the data into two homogeneous groups with respect to overall survival [10].

We expected Faraggi's method to be efficient because the authors, in their simulation, showed that their method was almost unbiased when the relative risk was less than 1.5 and provided an underestimation of only 5% when the relative risk was greater than 1.5. This method also gave a satisfactory type I error under the null hypothesis and had a good power for a large relative risk, as expected for two different prognostic groups.

To our knowledge, this methodology has never been used for cut-off determination in quality of life studies.

b) To evaluate how these optimally selected QoL scales as well as other clinical factors could be used to improve the performance of prognostic systems.

Patients and methods

Patients

This prognostic study was conducted in parallel to the CHOC trial. The CHOC trial included 271 patients with HCC in

palliative setting between July 2002 and October 2003 from 79 centres in France. The phase III CHOC trial was designed to demonstrate the efficacy of long-acting octreotide for the treatment of advanced hepatocellular carcinoma. The negative results of this trial have been previously published [23]. The protocol was reviewed and approved by the Ethics Review Committee of Région Picardie, France (16th May 2002). All patients provided written informed consent and the study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki. QoL data and associated patient characteristics are detailed in a previous publication [3].

Health-related Quality of life tool

QoL was self-completed by the patient during the two weeks prior to randomization using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30) questionnaire [24, 3]. Details for the European Organization for Research and Treatment of Cancer QLQ-C30 scales and their scoring can be found in our previous publication [3]. The response for each scale of a dimension was transformed into a score between 0 and 100 [25].

The present study focused on six QLQ-C30 scales: global health, physical functioning, role functioning, fatigue, dyspnea and diarrhea because these scales were the most significantly associated with OS using univariate as well as multivariate Cox models [3]. For fatigue, dyspnea and diarrhea, 100 was the worst score whereas for global health, physical functioning and role functioning, 100 was the best score.

Definition of the Prognostic classification

The four prognostic classifications used in the present study (CLIP, BCLC, GRETCH and BoBar) are defined in Table 1.

Statistical methods

Methods for descriptive statistics of our population were described in our previous publication [3]. Overall survival was defined as the time from randomization to death (regardless of the cause) or last follow-up (censored data).

All randomized patients with complete baseline QoL data were included in the statistical analysis and constituted a modified intention-to-treat population.

CHILD PUGH	Scores	
	0	1
Presence of ascites	No	Yes
Tumor size (>50%)	No	Yes
Bilirubin (>50 µmol/L)	No	Yes
Albumin (>30g/L)	No	Yes

CLIP	Scores		
	0	1	2
Child pugh	A	B	C
Tumor morphology	Uninodular and extension ≤50%	multiinodular and extension ≤50%	Massive or extension >50%
AFP (>400 ng/d)	No	Yes	
Portal vein thrombosis	No	Yes	

BCLC	Scores						
	A1	A2	A3	A4	B	C	D
WHO PS	0	0	0	0	0	1-2	3-4
Tumor stage	Single	Single	Single	3 tumors<3 cm	Multinodular	Vascular invasion or extrahepatic spread	Any
Okuda	I	I	I	I-II	I-II	I-II	III
Liver functional status	No portal hypertension and normal bilirubin	Portal hypertension and normal bilirubin	Portal hypertension and abnormal bilirubin	Child-Pugh A-B	Child-Pugh A-B	Child-Pugh A-B	Child-Pugh C

GRETCH	Scores			
	0	1	2	3
Bilirubin (≥50µmol/l)	No			Yes
Alkaline-phosphatase (≥2N*)	No		Yes	
AFP (≥35µg/l)	No		Yes	
Portal vein thrombosis	No	Yes		
Karnofsky (<80%)	No			Yes

BoBar	Scores		
	0	1	2
Non-small HCC	No	Yes	
Portal vein thrombosis	No	Yes	
Metastasis	No	Yes	
WHO PS	0	1	2-3
Jaundice	No	Yes	
Ascites	No	Yes	
AFP (>200 µg/L)	No	Yes	
Alkaline phosphatase (>2N *)	No		Yes

CLIP: Stage I (0); stage II (1-2); stage III (3-4-5)
BCLC: stage A to D
GRETCH: A (0) B (1-5) C (6-11)
BoBar: Low risk (0-3); intermediate risk; (4-6) high risk (7-10)
WHO PS= World Health Organization's performance; AFP= alpha-fetoprotein; * = More than twice the upper limit of normal
Table1. Definition of prognostic classifications.

Before describing the methods used to determine cut-points, the set of potential cut-points for each QoL scale was selected as follows [9]:

- First, all possible values of the QoL scale were selected.
- Values of the QoL scale below and above its 20^{th} and 80^{th} percentile, respectively, were removed to avoid a marked disequilibrium between the two

groups.

- For each value θ of the QoL scale between the 20th and the 80th percentile, the shortest time at which 80% of the patients had died ($T_{80}(\theta)$) was determined.
- The $\theta/T_{80}(\theta)$ curve was plotted.
- The monotonicity of the curve was studied and relevant cut-points were selected by avoiding redundant cut-points corresponding to a constant portion of the curve.

Faraggi's method [19] is a two-fold cross-validation consisting of partitioning the overall sample into learning and validation sub-samples. An optimal cut-point is determined for each sub-sample by using the minimum p-value approach (the value associated with the maximal log-rank statistic or equivalently the minimum p-value) and each patient was classified according to the cut-point of the sub-sample to which the patient did not belong. The final cut-point was the value (among all possible cut-points) that minimized the p-value in the overall sample using a stratified log-rank test with sub-sample as the stratum. Stability of the cut-points was studied with 500 bootstrap replications. The recommended cut-point was the most frequent one across 500 bootstrap replications [26]. Confidence intervals for cut-points were based on percentiles of the distribution.

Once an optimal cut-point was selected, the log-hazard ratio and its 95% confidence interval (95%CI) were computed using the method described by Höllander [27].

In view of the possible multicollinearity problem for global health reported by Van Steen [28] and the difficulty of resolving this problem (in the case of prognostic value) to improve outcome, only the other five dichotomized QoL scales as well as clinical and laboratory variables were selected for improvement of the prognostic classification using a multivariable Cox proportional hazards model with a backward elimination procedure (the prognostic classification was imposed in the model). Improvement of the performance of prognostic classifications was evaluated by Harrell's C-index [20], category-less net reclassification improvement (NRI) [22] and integrated discrimination improvement [21] (IDI). The last two statistics were computed at 3, 6 and 12 months. As stated by d'Agostino et al. [22], NRI "quantifies the correctness of upward and downward reclassification or movement of predicted probabilities as a result of adding a new marker", IDI quantifies the improvement in the sensitivity to predict mortality (without sacrificing specificity), whereas C-index

evaluates the discriminative ability of a model and ranges from 0.5 (no discrimination) to 1 (perfect discrimination).

The construction of a modified prognostic index was based on linear transformation (the regression coefficients were divided by the smallest one) and patients were arbitrary divided into three risk groups.

Survival curves were constructed using the Kaplan-Meier method [29].

All statistical analyses were carried out using the open-source R.2.12.0 software. IDI and NRI were computed using the survIDINRI library.

Results

Patient characteristics

Baseline QoL scores for the six scales were available for 214 (79%) of the 271 patients. More information about patient characteristics is available in our previous study [3].

Cut-point definition

Role functioning:

The median role functioning score was 83 (range: [0 – 100]). Figure 1A shows the Kaplan-Meier predictive failure time at which 80% of patients died as a function of role functioning scale. The choice of potential cut-points was as follows: 50 and 66.66. Results of optimal cut-points and the corresponding hazard ratio are summarized in Table 2. The most frequently selected cut-point across 500 bootstrap replications (418/500BRs) was 66.66 (95%CI=[50, 66.66]) and the rank of log-hazard ratios is shown in Figure 1C. The learning and validation sub-samples found the same cut-points 326 times out of 500 bootstrap replications (326/500BRs) as shown in Figure 1B. The corresponding hazard ratio was 1.76 (95%CI=[1.31 – 2.36]).

Results for cut-off determination for the other five QLQ-C30 scales are summarized in Table 2 and Figure 1B.

Revised prognostic classifications

Performances of the various prognostic classifications are shown in Table 3.

This section describes how the optimally dichotomized QoL scales and other clinical factors can be added to well-established prognostic systems. Modified prognostic indices are presented in Table 4.

QLQ-C30 QoL scales	Cutoff 95%CI	frequency	HR (95%CI)	Frequency of identical cutoff between learning and validation sample
Global health	50[50- 83.33]	226/500	1.61 [1.39 – 1.87] (<50 vs. ≥50)	19/500
Physical functioning	58.33 [58.33 – 66.66]	436/500	1.51 [1.42 – 1.59] (<58.33 vs. ≥58.33)	355/500
Role functioning	66.66[50 – 66.66]	418/500	1.76 [1.31 – 2.36] (<66.66 vs. ≥66.66)	326/500
Fatigue	66.66 [16.66 – 66.66]	303/500	2.09 [1.83 – 2.39] (>66.66 vs. ≤66.66)	116/500
Diarrhoea	33.33 [0 – 33.33]	356/500	1.62 [1.38 – 1.90] (>33.33 vs. ≤33.33)	188/500
Dyspnoea	0[0 – 66.66]	295/500	1.48 [1.27 – 1.73] (>0 vs. 0)	139/500

Table 2: Results and frequency of optimal cut-points determination using the Faraggi method.

	C-index (95% CI)	NRI(95% CI)	IDI(95% CI)
Crude pronostic classifications			
BCLC	0.57 [0.53 - 0.60]		
CLIP	0.62 [0.59 – 0.65]		
BOBAR	0.63 [0.60 – 0.67]		
GRETCH	0.59 [0.56 - 0.62]		
Improvement of prognostic classifications with continuous variables			
BCLC Oedema (2.47 [1.58 – 3.87]) Portal vein thrombosis (2.06 [1.38 – 3.08]) Alpha-fetoprotein (1.88 [1.40 – 2.52]) Alkaline phosphatase (1.56 [1.11 – 2.21]) fatigue (1.008 [1.003 – 1.013]) Diarrhoea (1.006 [1.00 – 1.011])	0.68 ([0.65 – 0.72])	0.98 [0.60 – 1.20] (3 months) 0.64 [0.32 – 0.94] (6 months) 0.60 [0.26 – 0.92] (12months)	0.15 [0.08 – 0.22] (3 months) 0.15 [0.08 – 0.22] (6 months) 0.14 [0.07 – 0.20] (12 months)
CLIP Oedema 2.02 [1.34 – 3.04] Hepatomegaly 1.49 ([1.01 – 2.04] Fatigue 1.010 [1.005 – 1.015] Diarrhoea 1.007 [1.001 – 1.012]	0.68 ([0.64 – 0.72])	0.64 [0.20 1.02] (3 months) 0.40 [-0.02 – 0.68] (6 months) 0.40 [-0.02 – 0.60] (12 months)	0.09 [0.03 – 0.15] (3 months) 0.08 [0.02 – 0.12] (6 months) 0.05 [0.01 – 0.10] (12 months)
BOBAR Dyspnoea 1.005 [1.001 – 1.010] Diarrhoea 1.007 [1.001 – 1.012]	0.67 ([0.63 – 0.71])	0.32 [-0.06 – 0.70] (3 months) 0.14 [-0.12 – 0.44] (6 months) 0.26 [0.00 – 0.64] (12 months)	0.02 [-0.01 – 0.07] (3 months) 0.02 [-0.01 – 0.05] (6 months) 0.03 [0.00 – 0.06] (12 months)
GRETCH Oedema (2.18 [1.45 – 3.28]) fatique (1.008 [1.003 – 1.013]) Diarrhoea (1.007 [1.001 – 1.013])	0.66 ([0.62 – 0.70])	0.64 [0.30 – 0.98] (3 months) 0.40 [0.00 – 0.74] (6 months) 0.40 [0.00 – 0.64] (12 months)	0.09 [0.04 – 0.16] (3 months) 0.08 [0.02 – 0.14] (6 months) 0.05 [0.01 – 0.10] (12 months)
Improvement of prognostic classifications optimally dichotomized QoL items and other clinical variables.			
BCLC Oedema 2.31 [1.49 – 3.59] Portal vein thrombosis 2.00 [1.34 – 2.98] AFP 1.94 [1.45 – 2.59] Alkaline phosphatase 1.62 [1.15 – 2.28] Fatigue 1.86 [1.25 – 2.77] Diarrhoea 1.67 [1.04 – 2.68]	0.66 [0.62 – 0.69]	0.78 [0.40 – 1.06] (3 months) 0.54 [0.28 – 0.81] (6 months) 0.62 [0.28 – 0.86] (12 months)	0.10 [0.05 – 0.15] (3 months) 0.12 [0.04 – 0.19] (6 months) 0.14 [0.06 – 0.23] (12 months)
CLIP Oedema 1.76 [1.18 – 2.63] Hepatomegaly 1.44 [1.06 – 1.94] physical functioning 1 .49 [1.03 – 2.17] Fatigue 2.09 [1.39 – 3.13]	0.65 [0.61 – 0.69]	**Compared to the original CLIP** 0.58 [0.28 – 0.86] (3 months) 0.32 [0.00 – 0.58] (6 months) 0.30 [0.10 – 0.52] (12 months) **Compared to CLIP+PS** 0.26 [0.00 – 0.84] (3 months) -0.04 [-0.28 – 0.48] (6 months) -0.12 [-0.32 – 0.54] (12 months)	**Compared to the original CLIP** 0.06 [0.01 – 0.11] (3 months) 0.04 [-0.01 – 0.10] (6 months) 0.04 [0.00 – 0.09] (12 months) **Compared to CLIP+PS** 0.06 [0.01 – 0.11] (3 months) 0.04 [-0.01 – 0.09] (6 months) 0.03 [-0.01 – 0.08] (12 months)
BOBAR Fatigue 1.86 [1.28 – 2.71]	0.65 [0.61 – 0.68]	0.24 [-0.48 – 0.54] (3 months) 0.12 [-0.64 – 0.36] (6 months) 0.22 [-0.88 – 0.48] (12 months)	0.02 [-0.01 – 0.06] (3 months) 0.01 [-0.03 – 0.04] (6 months) 0.02 [-0.02 – 0.05] (12months)
GRETCH Oedema 2.03 [1.36 – 3.03] Fatigue 1.79 [1.19 – 2.70] Diarrhoea 1.80 [1.10 – 2.93]	0.64 ([0.60 – 0.67])	0.58 [0.24 – 1.06] (3 months) 0.28 [0.02 – 0.76] (6 months) 0.38 [0.12 – 0.72] (12 months)	0.06 [0.02 – 0.13] (3 months) 0.05 [0.00 – 0.12] (6 months) 0.07 [0.01 – 0.11] (12 months)

Table 3: Performances of the prognostic classifications with QoL scales treated as continous or dichotomized variables.

BCLC

The C-index was 0.57 (95%CI=[0.53 - 0.60]) for the BCLC score with four categories. As categories A and B tended to have a similar survival in our palliative population, these two categories were pooled before analyzing the improvement of performance; after pooling, the C-index remained unchanged (0.57 (95%CI=[0.53 - 0.60])). The modified BCLC score is defined in Table 4A with a 9% increase in C-index (from 0.57 to 0.66 (95%CI=[0.62 – 0.69])); this gain in discrimination of the prognostic groups is illustrated in Figures 2A&2B. The 3-month NRI and IDI were 0.78 (95%CI=[0.40 – 1.06]) and 0.10 (95%CI=[0.05 – 0.15]), respectively. Results for 6-month and 12-month NRI and IDI are summarized in Table 3.

CLIP

The new CLIP score integrating optimally dichotomized QoL score is defined in Table 4B with a 3% gain for C-index (from 0.62 for original CLIP to 0.65 [0.61 – 0.69]). The improvement in separation of prognostic groups was also assessed in terms of NRIs and IDIs (Table 3), as illustrated in Figures 3A&3B.

BoBar

The improvement of performance for the modified BoBar (Table 4C) was limited in terms of C-index (2%), NRIs and IDIs (Table 3), as illustrated by supplementary Figures 1A&1B.

GRETCH

An absolute 5% improvement of C-index (from 0.59 to 0.64 [0.60 – 0.67]) was observed for the modified GRETCH (Table 4D) prognostic system compared to GRETCH alone.

Survival curves for the original and modified GRETCH are shown in supplementary Figures 2A&2B.

Discussion

This study established cut-points for the six most important QoL scales in terms of overall survival prognosis.

Patients could be divided into two homogeneous prognostic groups using cut-points 50, 66.66, 58.33, 66.66, 0 and 33.33 for global health, role functioning, physical functioning, fatigue, dyspnea and diarrhea, respectively.

	0 point	1 point	2 points	3 points	4 points	5 points	6 points	7 points	8 points
Table 4A: MODIFIED BCLC									
BCLC	BCLC class A/B	BCLC class C	BCLC class D						
Oedema	No			Yes					
Portal vein thrombosis	No		Yes						
Alpha-fetoprotein	<=200		>200						
Alkaline phosphatase	<N	>2N							
Fatigue	<=66.67	>66.67							
Diarrhoea	<=33.33	>33.33							
Table 4B: MODIFIED CLIP									
CLIP	CLIP I	CLIP II							CLIP III
Oedema	No				Yes				
Hepatomegaly	absence		presence						
Fatigue	<=66.67					>66.67			
physical functioning	>=58.33			<58.33					
Table 4C: MODIFIED BoBar :									
BoBar	class I	class II	class III						
Fatigue	<=66.67	>66.67							
Table 4D: MODIFIED GRETCH									
GRETCH	class A	class B		class C					
Oedema	No	Yes							
Fatigue	<=66.67	>66.67							
Diarrhoea	<=33.33	>33.33							

The modified BCLC varied from 0 to 12.
Poor prognosis: 4-13; Intermediate prognosis: 2-3; Good prognosis: 0-1
The modified CLIP varied from 0 to 22.
Poor prognosis: 12-22; Intermediate prognosis: 8-11; Good prognosis: 0-7
The modified BoBar varied from 0 to 3.
Poor prognosis: 2-3; Intermediate prognosis: 1; Good prognosis: 0
The modified GRETCH varied from 0 to 6.
Poor prognosis: 2-6; Intermediate prognosis: 1; Good prognosis: 0

Table4: Definition of the four revised prognostic classifications.

Although the extremely low cutoff value for dyspnea may be related to the low frequency of dyspnea symptoms, our results suggested that the presence of dyspnea has a prognostic significance and should be closely monitored by the medical team and appropriate actions should be taken according to the severity of the dyspnea. This finding meant that any moderate or severe dyspnea should be managed appropriately in patients with advanced hepatocellular carcinoma, although dyspnea was identified as a prognostic factor on multivariate Cox analysis, suggesting that the presence of dyspnea may be a consequence of a clinical or laboratory factor already present in the prognostic classification.

These cut-points can be easily used to define eligibility criteria, stratification factors or as binary endpoints for future trials including palliative HCC patients.

The four prognostic systems most commonly used for HCC patients could be revised by using these optimally selected cut-points for QoL scales. Almost all revised prognostic classifications clearly improved the accuracy of overall survival prediction and each classification (except BoBar) included two QoL scales. Moreover, the variables added to each prognostic classification after dichotomization of QoL scales were very similar to those added when QoL scales were treated as continuous variables. Furthermore, Harrell's C-indices did not vary substantially for all the prognostic systems regardless of the type of QoL scale analysis (continuous vs. dichotomized; table 3); our proposed cut-points can therefore be considered as optimal in terms of prognosis.

On average, IDI was significantly different from zero, indicating that inclusion of QoL scales in the prognostic classification was associated with a more marked improvement of the sensitivity to detect patients likely to die by a defined time-point. IDI was uniformly good for the revised BCLC and CLIP staging systems (compared to the original BCLC and CLIP respectively). On average, a greater than 10% improvement was observed for the sensitivity to predict death for the new BCLC, regardless of the time-point (three, six or twelve months) whereas the improvement in sensitivity was about 5% for the new CLIP. IDI was significantly different from zero (95%CI did not contain zero) at three months but not at six and twelve months for the revised CLIP (compared to CLIP+World Health Organization performance status). This result highlighted the fact that taking QoL scales into account could improve identification of patients likely to die within three months compared to the CLIP+ World Health Organization performance status. In other words, QoL scales assessed by the patient allow more accurate detection of the patient's symptom burden than the World Health Organization performance status completed by the clinician. This result is concordant with the findings of Efficace et al. [30] about physicians' underestimation of symptoms for patients with chronic myeloid leukemia. Quality of life scales could therefore be selected as inclusion/exclusion criteria as well as stratification factors. These staging systems including QoL scales allow the physician to take into account the patient's perception of his/her disease.

The good performance of the revised BCLC was achieved after adding six new variables, confirming the limited prognostic value of BCLC alone in advanced HCC patients [16, 17].

We did not perform a sensitivity analysis because previous results [3] showed that imputing missing QoL scales did not significantly change the results of complete case analysis.

Recently, the European Association for the Study of the Liver - European Organization for Research and Treatment of Cancer (EASL-EORTC) group [17] stated that, in cancer studies, QoL is the third most important endpoint in terms of strength of evidence after overall mortality and cause-specific mortality. QoL was then ranked above the well-known surrogate endpoints in oncology: progression-free survival, disease-free survival, time to treatment failure and tumor response; these endpoints are defined with a binary event (presence vs. absence). A binary definition of QoL, as proposed in this study, could therefore facilitate the definition of a recommended target value for a given QoL scale and its use by investigators as an endpoint for phase II as well as phase III trials.

The widely accepted European Organization for Research and Treatment of Cancer QLQ-C30 questionnaire translated into several languages and used in numerous clinical trials in oncology was used in this study; however, unfortunately, our study did not include the HCC18 module for HCC patients which constituted a limitation. It would have been preferable to identify the independent prognostic components of this module and to propose cut-points for each component and to include them in the revised prognostic systems. Although internal validity by bootstrapping was good, our revised prognostic indices need to be prospectively validated in independent cohorts of advanced HCC patients.

From a statistical point of view, Faraggi and Simon [19] showed by simulation that their cross-validation method controls the type I error, thus avoiding the frequently reported inflation of type I error in studies designed to detect an optimal cut-point.

We believe that our study addresses the need to refine the original BCLC class-C [16, 17] corresponding to patients with advanced HCC and that our revised BCLC, CLIP and BoBar staging systems could be used for these patients.

Our work constitutes a step in the direction of the recommendations proposed by Gotay et al. [31] to determine the appropriate scales and cut-points for stratification and eligibility determination. We expect that similar research will be performed in other types of cancer.

We believe that this dichotomization of QoL scales will facilitate integration of QoL scales in decision-making for the treatment of advanced HCC patients as well as their use in future clinical trial planning.

Conclusion: The cutoff points for the six QoL scales could be used to evaluate the well-being of patients with advanced HCC before starting any treatment, as most patients receive Sorafenib for which the most common adverse effects are diarrhea and fatigue. Patients with these two symptoms before treatment should therefore be closely monitored. The cutoff points could also be used alone (or in revised prognostic classifications after prospective validation) in the design of clinical trials to defined eligibility criteria or stratification factors.

Abbreviations: QoL, quality of life; HCC, hepatocellular carcinoma; CLIP, Cancer of the Liver Italian Program; BCLC, Barcelona Clinic Liver Cancer; GRETCH, Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire; BoBar, BOnnetain & BARbare prognostic index; NRI, net reclassification improvement; IDI, integrated discrimination improvement;

Competing interests:

None declared.

Acknowledgments: We thank the FFCD (Fédération Française de Cancérologie Digestive) for data supply. We also thank the CRCI (Coordination de la Recherche Clinique et de l'Innovation, CHU Amiens) and RICH (Réseau des Investigateurs pour le Carcinome Hépatocéllulaire) for funding sources (correction of English language).
We thank the reviewer for his/her useful comments.

References

1. A. Jemal, Center, M M, C. DeSantis, E. M. Ward. Global Patterns of Cancer Incidence and Mortality Rates and Trends. Cancer Epidemiology Biomarkers & Prevention. 2010; 19(8); 1893-1907.
2. Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. Lancet. 2003;362(9399):1907-17.
3. Diouf M, Filleron T, Barbare JC et al. The added value of quality of life (QoL) for prognoses of overall survival in patients with palliative hepatocellular carcinoma. J Hepatol. 2013; 58: 509-521.

4. Bonnetain F, Paoletti X, Collette S et al. Quality of life as a prognostic factor of overall survival in patients with advanced hepatocellular carcinoma: results from two French clinical trials. Qual Life Res 2008; 17(6):831-43.
5. Yeo W, Mo F.K.F, Koh J et al. Quality of life is predictive of survival in patients with unresectable hepatocellular carcinoma. Ann Oncol 2006; 17: 1083–1089
6. Quinten C, Coens C, Mauer M et al. Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. Lancet Oncol 2009; 10:865-71.
7. Royston P, Altman DG, Sauerbrei W (2006). Dichotomizing continuous predictors in multiple regression: a bad idea. Statist. Med 25:127–141.
8. Roychowdhury DF, Hayden A and Liepa AM. Health-related quality-of-life parameters as independent prognostic factors in advanced or metastatic bladder cancer. J Clin Oncol. 2003;21(4):673-8.
9. Mazumdar M and Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and application to decision-making about cancer treatments. Statist. Med 2000; 19, 113–132.
10. Williams BA, Mandrekar JN, Mandrekar SJ et al. Finding optimal cutpoints for continuous covariates with binary and time-to-event outcomes. 2006.
11. Wartenberg D, Northridge M. Defining exposure in case-control studies: a new approach. American Journal of Epidemiology 1991; 133:1058 –1071.
12. Altman DG, Lausen B, Sauerbrei W, Schumacher M. The dangers of using ‘optimal’ cutpoints in the evaluation of prognostic factors. Journal of the National Cancer Institute 1994; 86:829–835.
13. The Cancer of Liver Italian Program Investigators. A new prognostic system for hepatocellular carcinoma: A retrospective study of 435 patients: The Cancer of the Liver Italian Program (CLIP) investigators. J Hepatol 1998; 28:751–755.
14. Llovet JM, Bru C and Bruix J. Prognosis of hepatocellular carcinoma: The BCLC staging classification. *Seminars in Liver Disease* 1999; 19, 329–338.
15. Chevret S, Trinchet JC, Mathieu D et al.. A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d’Etude et de Traitement du Carcinome Hepatocellulaire. J Hepatol 1999; 31:133–141.
16. Tournoux-Facon C, Paoletti X, Barbare JC et al. Development and validation

of a new prognostic score of death for patients with hepatocellular carcinoma in palliative setting. J Hepatol 2011; 54(1):108-14.

17. EASL–EORTC Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol. 2012; 56: 908–943.
18. Huitzil-Melendez FD, Capanu M, M. O'Reilly E et al. Advanced Hepatocellular Carcinoma: Which Staging Systems Best Predict Prognosis? J Clin Oncol. 2010; 28:2889-2895
19. Faraggi D and Simon R. A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis. Statist. Med 1996; 15, 2203–2213.
20. Harrell FE, Lee KL and Mark DB. Tutorial in biostatistics. Multivariable prognostic models issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statist. Med 1996*;* 15:361-387.
21. Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008; 27:157-72.
22. Pencina MJ, Ralph B. D'Agostino, Ewout W. Steyerberg (2011). Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Statist. Med 2011*;* 30:11-21.
23. Barbare JC, Bouche O, Bonnetain F et al. Treatment of advanced hepatocellular carcinoma with long-acting octreotide: a phase III multicentre, randomised, double blind placebo-controlled study. *Eur J Cancer* 2009; 45(10):1788-97.
24. Aaronson NK, Ahmedzai S, Bergman B 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 1993; Inst.;85(5):365-76.
25. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group: The EORTC QLQ-C30 Scoring Manual 3rd edition. European Organization for Research and Treatment of Cancer, Brussels, Belgium; 2001.
26. Vinh-Hung V, Verkooijen HM, Fioretta G, et al. Lymph node ratio as an alternative to pN staging in node-positive breast cancer. J Clin Oncol

2009;27:1062-1068.

27. Höllander N, Sauerbrei W, Schumacher M. Confidence intervals for the effect of a prognostic factor after selection of an `optimal' cutpoint. Statistics in Medicine, 2004; 23:170-713
28. Van Steen K, Curran D, Kramer et al. Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. Stat Med. 2002; 21(24):3865-84
29. Kaplan EL, Meier P. *N*on parametric estimation from incomplete observations*.* J Am Stat Assoc 1958*;* 53:457-481
30. Efficace F, Rosti G, Aaronson N et al. Patient versus physician symptom and health status reporting in chronic myeloid leukemia. Haematologica. 2013; 98:xxx doi:10.3324/haematol.2013.093724
31. Gotay CC, Kawamoto CT, Bottomley A et al. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol 2008;*26(8):1355-1363.

Figure 1A

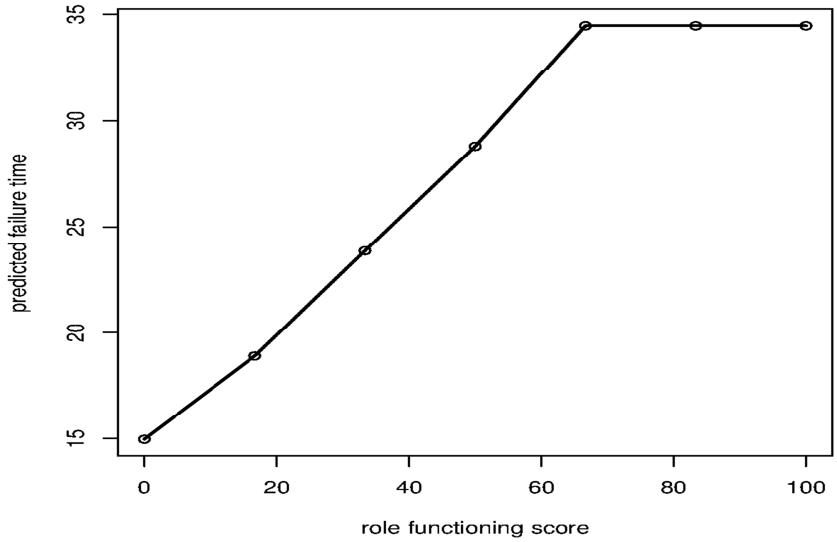
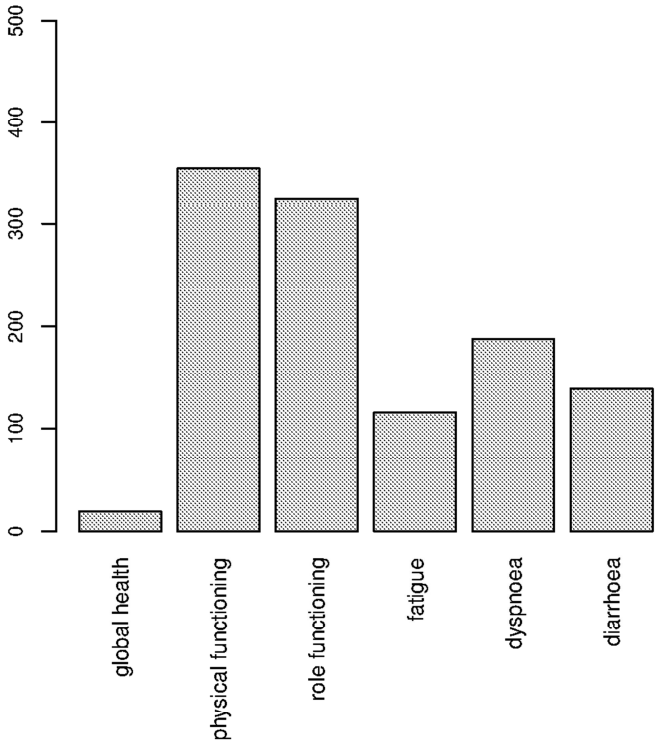

predicted failure time
15
20
25
30
35
0
20
40
60
80
100
role functioning score

Figure 1B


0
100
200
300
400
500
global health
physical functioning
role functioning
fatigue
dyspnoea
diarrhoea

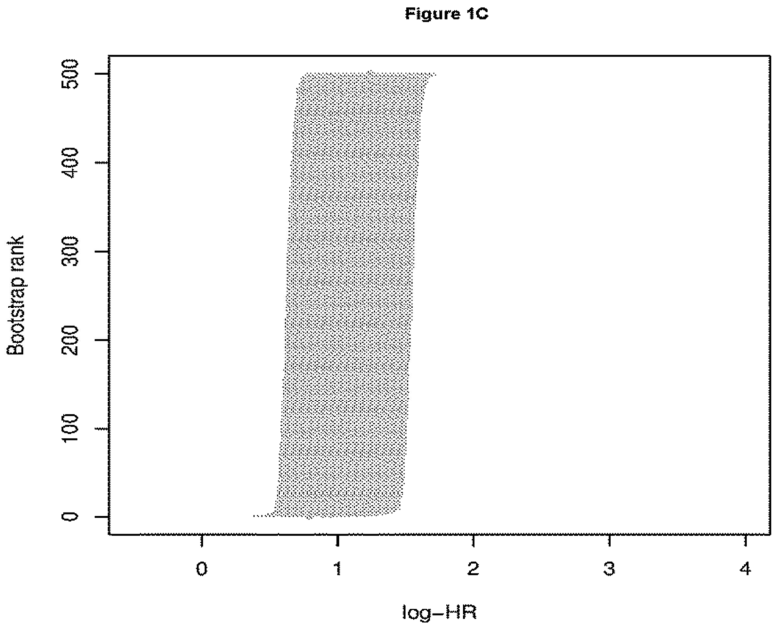


Figure 1: Results of determination of optimal cut-points for the QLQ-C30 role functioning scale.

A) Scatter plot of predictive time to observe 20% survivors as a function of role functioning score.

B) Frequency with which the learning and validation samples found the same cutoff for the six quality of life scales (N=500 bootstrap replications).

C) The log-hazard ratio and 95%CI ranks after 500 bootstrap replications for role functioning score.

Figure 2A

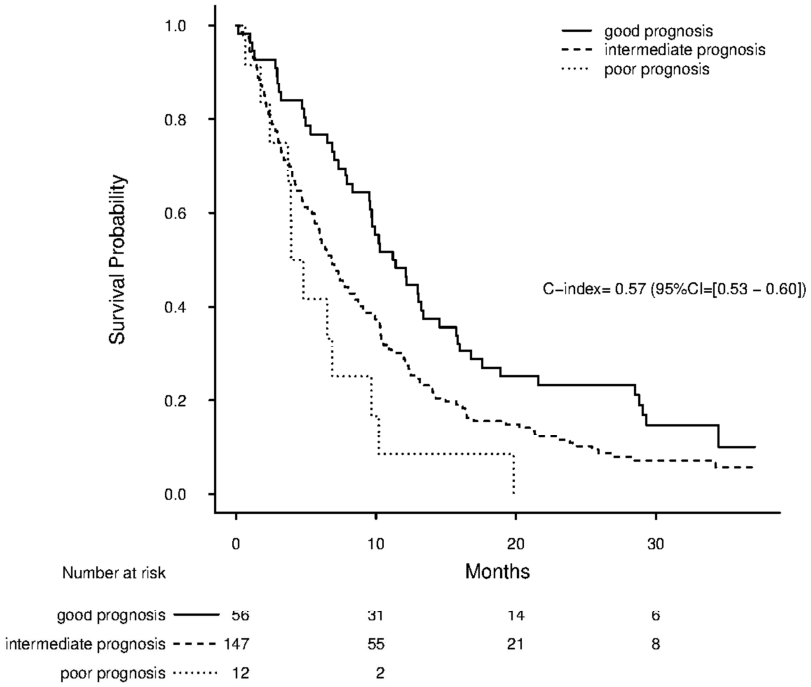

good prognosis
intermediate prognosis
poor prognosis
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
C–index= 0.57 (95%CI=[0.53 – 0.60])
0
10
20
30
Months
Number at risk
good prognosis 56 31 14 6
intermediate prognosis 147 55 21 8
poor prognosis 12 2

Figure 2B

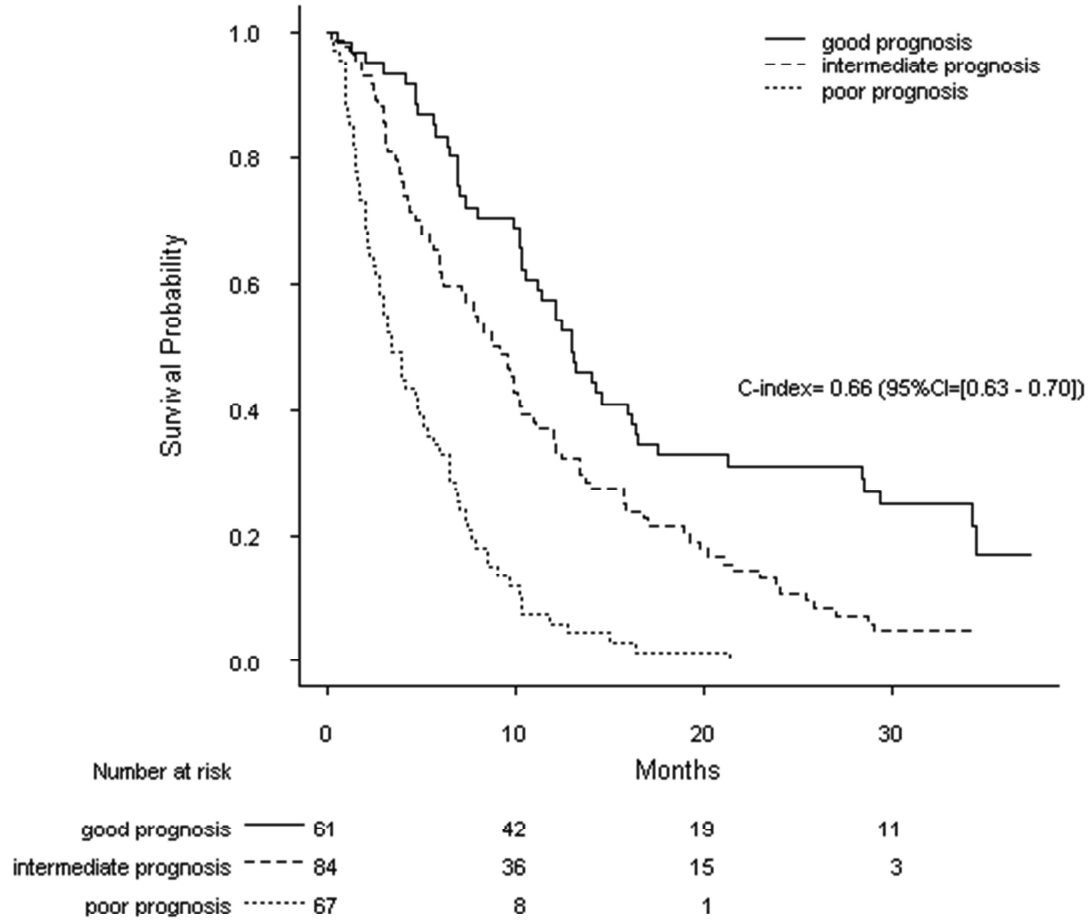


Figure 2: Overall survival (in months) for good, intermediate and poor prognosis according to the original and modified BCLC prognostic systems.

A) Median survival was 11.20 [9.57 – 15.73], 6.87 [5.77 – 8.73] and 3.93 [3.70 – NA] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively according to the original BCLC index.

B) Median survival was 13.03 [10.50 – 16.47], 8.77 [6.20 – 11.10] and 3.4 [2.8 – 5.33] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively according to the modified BCLC prognostic system.

Figure 3A

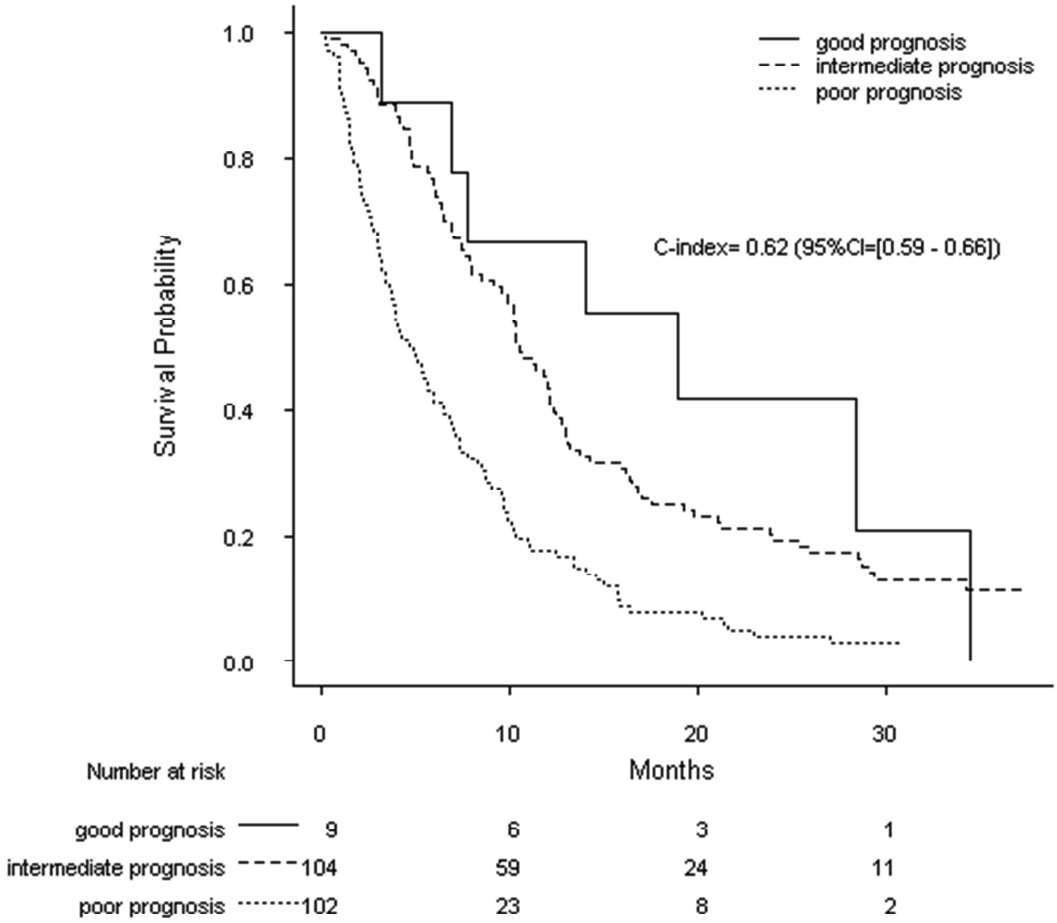

1.0
0.8
0.6
0.4
0.2
0.0
Survival Probability
good prognosis
intermediate prognosis
poor prognosis
C-index= 0.62 (95%CI=[0.59 - 0.66])
0
10
20
30
Months
Number at risk
good prognosis 9 6 3 1
intermediate prognosis 104 59 24 11
poor prognosis 102 23 8 2

Figure 3B

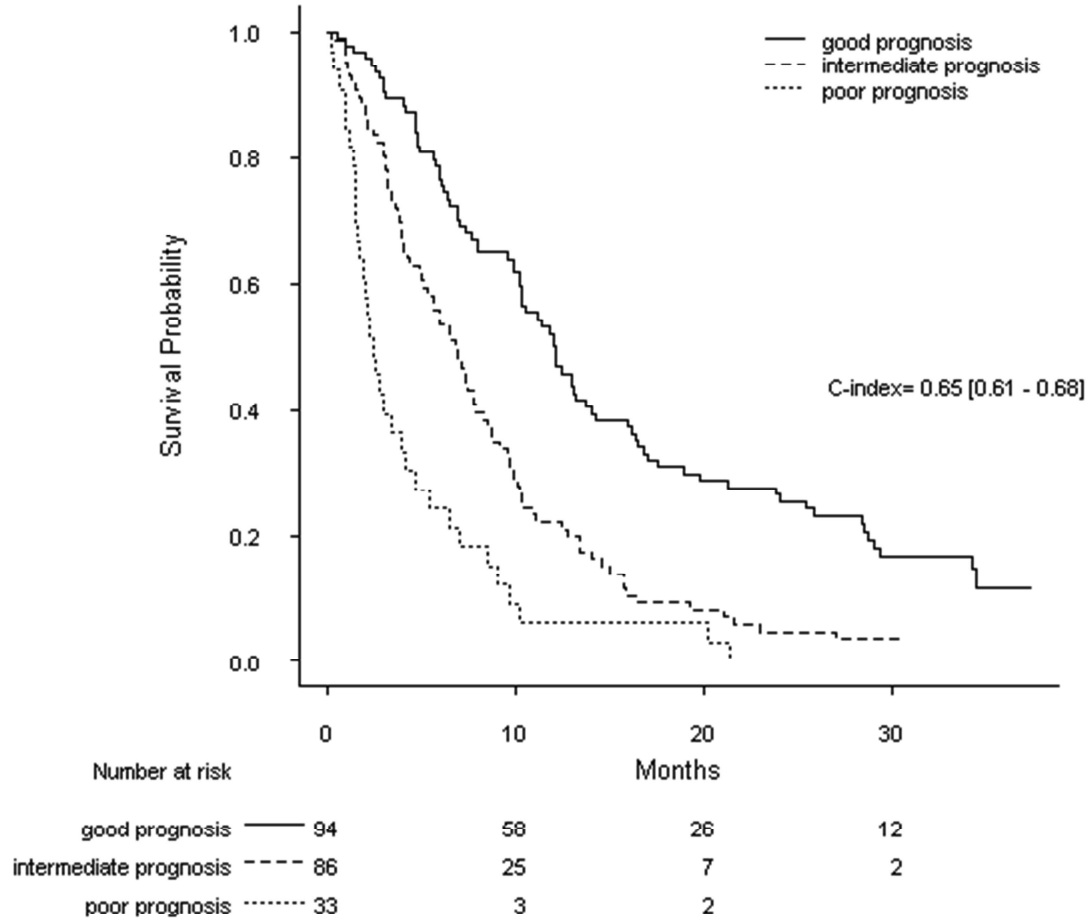
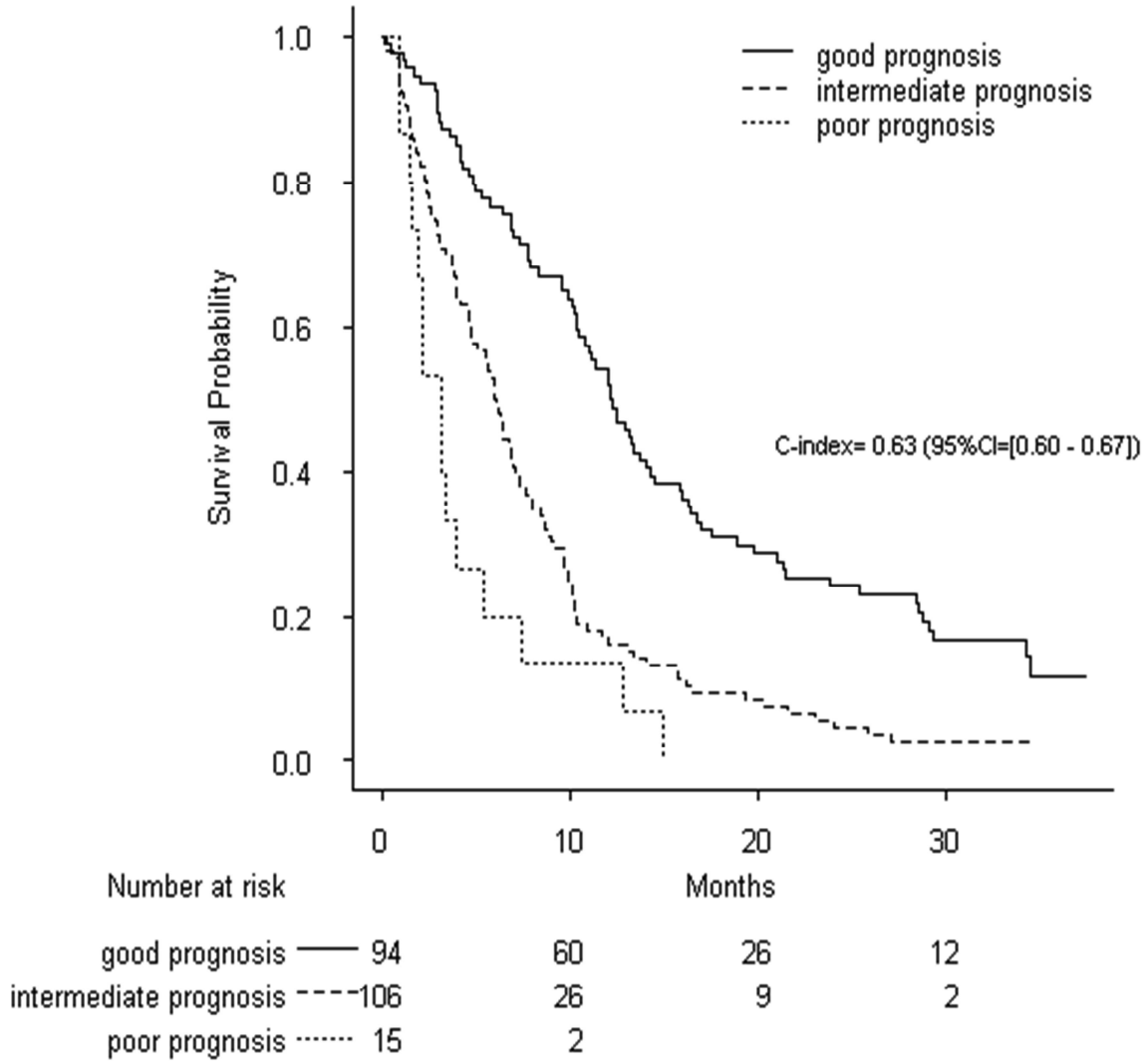


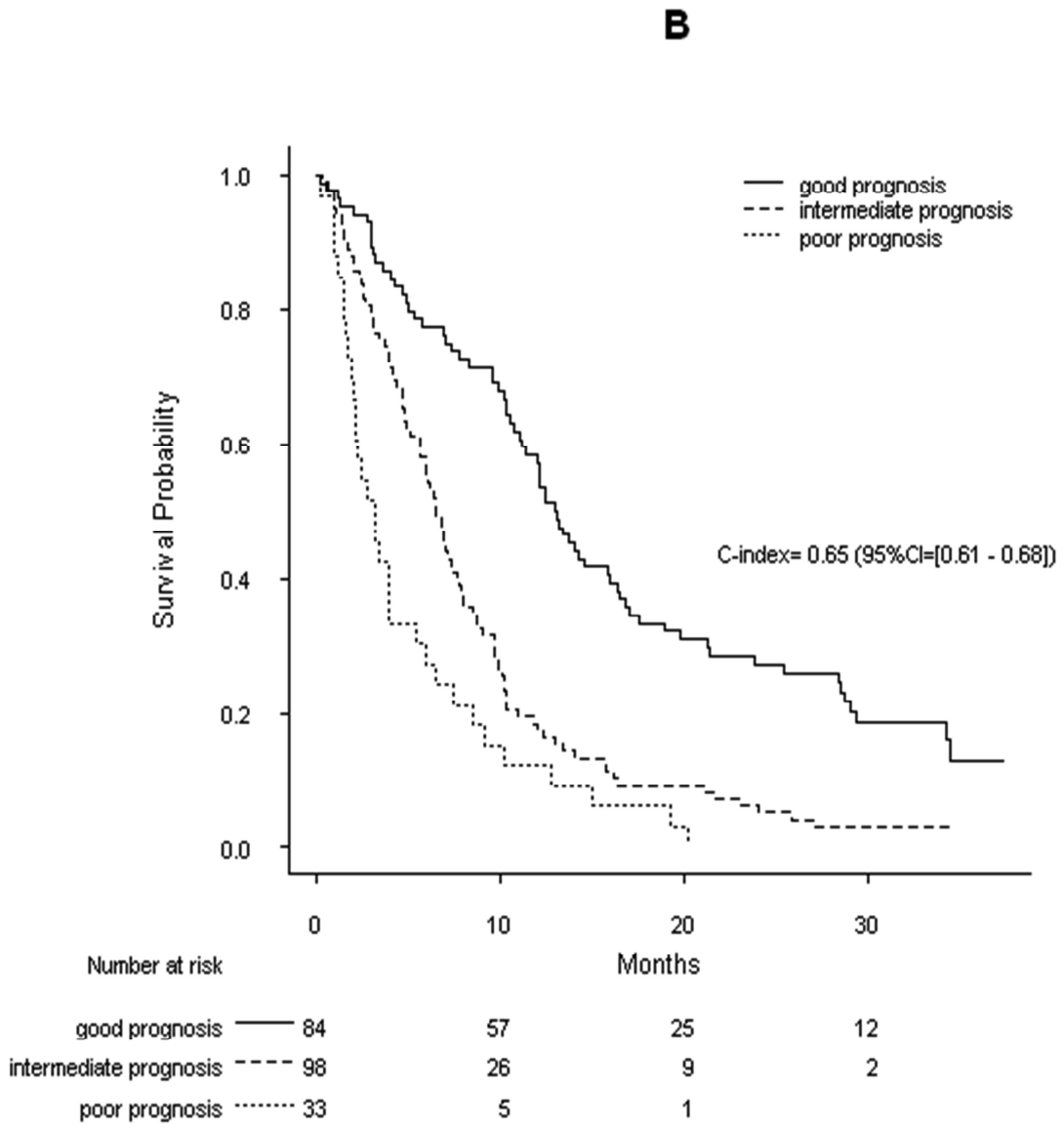
Figure 3: Overall survival (in months) for good, intermediate and poor prognosis according to the original and modified CLIP prognostic indices.

A) Median survival was 18.90 [7.77 – NA], 10.53 [9.57 – 12.50] and 4.77 [3.70 – 6.50] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively for the original CLIP index.

B) Median survival was 12.13 [10.30 – 16.40], 7.80 [5.63 – 1.30] and 3.93 [3.03 – 6.50] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively according to the modified CLIP index.

A

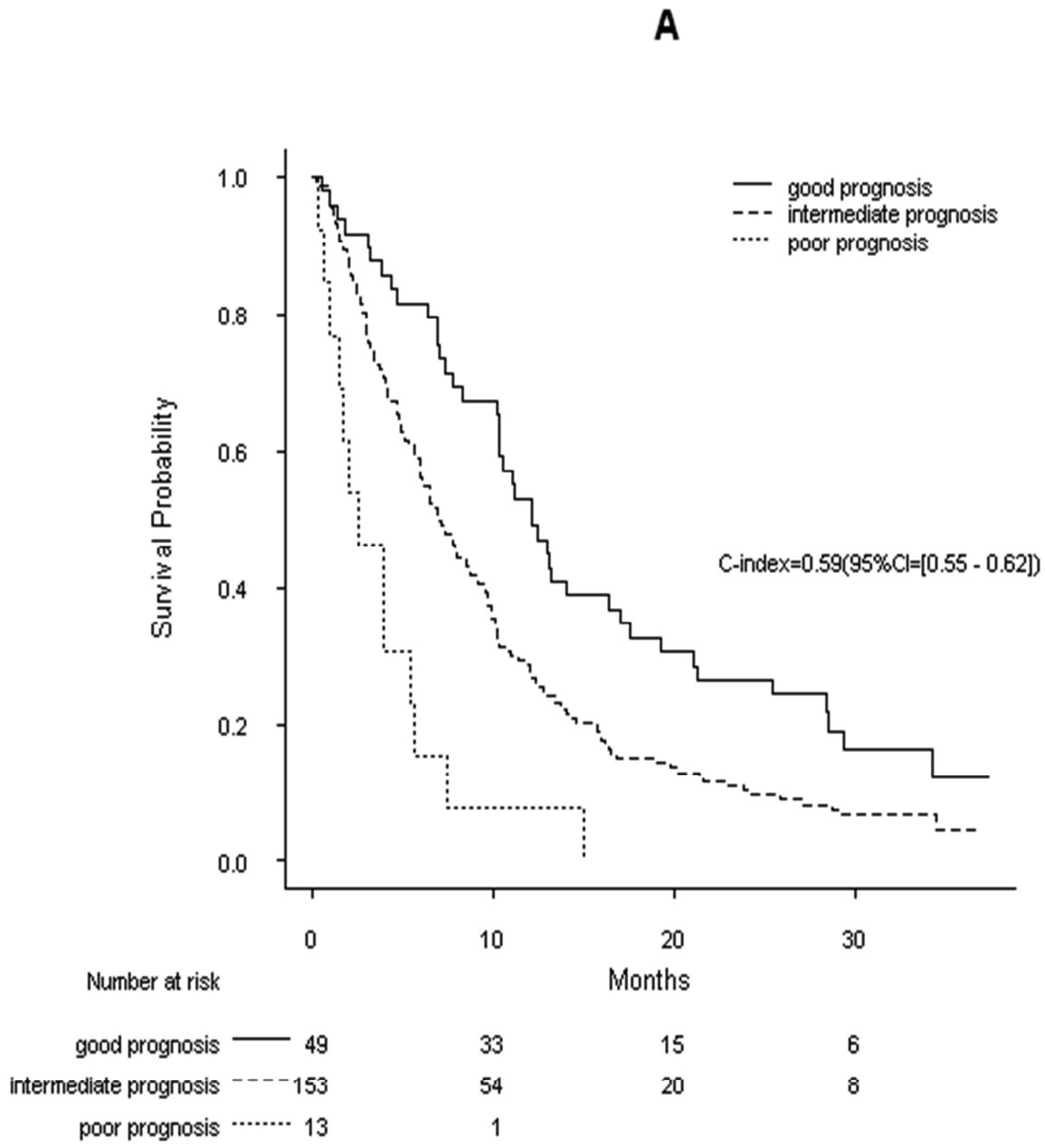

1.0
0.8
0.6
0.4
0.2
0.0
Survival Probability
good prognosis
intermediate prognosis
poor prognosis
C-index= 0.63 (95%CI=[0.60 - 0.67])
0
10
20
30
Months
Number at risk
good prognosis 94 60 26 12
intermediate prognosis 106 26 9 2
poor prognosis 15 2

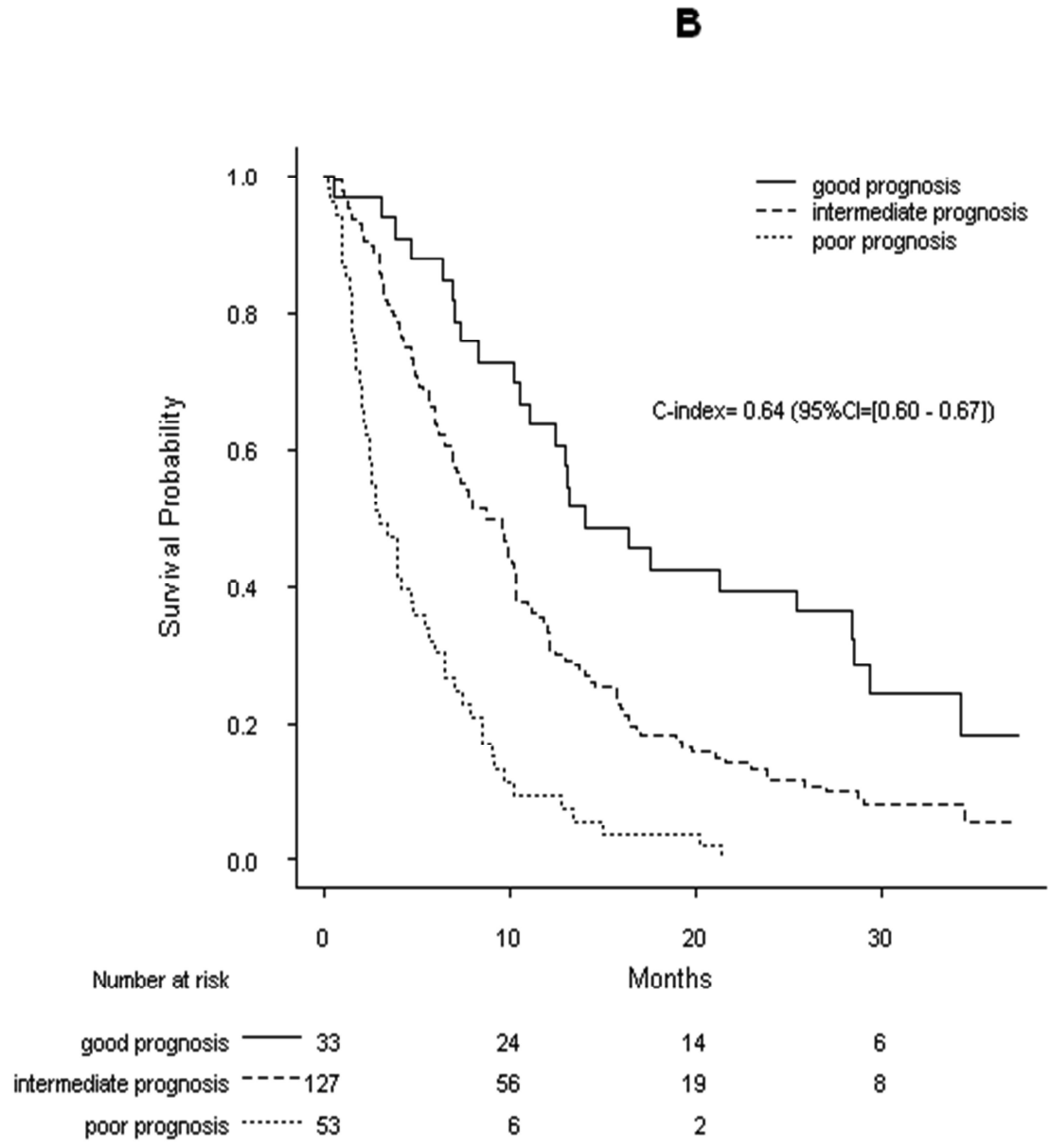


Supplementary figure 1: Overall survival (in months) for good, intermediate and poor prognosis according to the original and modified Bobar prognostic systems.

A) Median survival was 12.20 [10.50 – 14.53], 6.10 [4.80 – 7.30] and 3.20 [2.00 – 7.43] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively for the original Bobar system.

B) Median survival was 13.10 [11.20 – 16.47], 6.50 [5.63 – 7.77] and 3.20 [2.13 – 6.00] for the good (solid line), intermediate (dash line) and poor prognostic (dotted line) respectively for the modified Bobar system.


A
1.0
0.8
0.6
0.4
0.2
0.0
Survival Probability
good prognosis
intermediate prognosis
poor prognosis
C-index=0.59(95%CI=[0.55 - 0.62])
0
10
20
30
Months
Number at risk
good prognosis 49 33 15 6
intermediate prognosis 153 54 20 8
poor prognosis 13 1



Supplementary figure 2: Overall survival (in months) for good, intermediate and poor prognosis according to the original and modified GRETCH prognostic classifications

A) Median survival was 12.17 [10.30 – 17.60], 7.03 [6.00 – 9.07] and 2.57 [1.50 – NA] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively for the original GRETCH index.

B) Median survival was 14.10 [11.10 – 29.33], 8.77 [6.87 – 10.33] and 2.97 [2.40 – 5.43] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively for the modified GRETCH index.

iv. Adénocarcinome du pancréas : valeur pronostique de la qualité de vie.

1) Résumé :

Rationnel: L'IP-OMS a actuellement une place importante dans le choix du traitement pour les patients atteints d'adénocarcinome du pancréas (ACP). Cependant, plusieurs études dans le domaine du cancer ont montré la valeur pronostique de la qdv dans des populations homogènes selon l'IP-OMS.

Le but de cette étude a été d'identifier les scores de qdv ayant un intérêt pronostique pour les patients atteints d'ACP métastatique et d'établir des groupes homogènes de patients en fonction de leur pronostic.

Méthode: Les données provenant de 98 patients naïfs de toute chimiothérapie ayant un ACP prouvé histologiquement et recrutés entre 2007 et 2011 dans l'étude de phase II FIRGEM ont été analysées. L'étude FIRGEM était conçue pour montrer la supériorité d'une administration séquentielle de FOLFIRI.3/Gemcitabine par rapport à la Gemcitabine seule.

Les données de qdv ont été recueillies à l'aide du questionnaire QLQ-C30 de l'EORTC.

Une méthodologie de forêts aléatoires appliquées aux données de survie a été utilisée pour imputer les données manquantes et pour identifier les facteurs pronostiques importants.

Résultats: Les données de qdv ont été complétées par 60% des patients (59/98).

Quatorze variables pronostiques ont été identifiées. Les trois plus importants facteurs pronostiques étaient les scores de fatigue, d'activité quotidienne et de perte d'appétit suivis par l'ASAT et le CA19-9.

L'indice de discrimination de Harrell était de 0.65.

Conclusion: Les scores de qdv ont un intérêt dans le pronostic des patients atteints d'ACP métastatique. De plus, le score de fatigue, le degré d'autonomie dans les activités quotidiennes et la perte d'appétit ont montré une valeur pronostique plus importante que les paramètres clinico-biologiques.

Ces scores de qdv, particulièrement le score de fatigue, devrait faire partie du bilan clinique des patients atteints d'ACP métastatique ayant un bon état général.

2) Article sur l'adénocarcinome du pancréas:

Title: Prognostic value of quality of life in patients with metastatic pancreatic adenocarcinoma: a random forest methodology.

M. Diouf [1,2], T. Filleron[3], J Taieb[4], F. Bonnetain[2].

[1]Clinical Research and Innovation Directorate, Amiens University Hospital, Amiens, France.

[2]Methodology and quality of life in oncology Unit, EA 3181 CHU Besançon and the *Qualité de Vie et Cancer* clinical research platform, Dijon, France.

[3]Biostatistics Unit, Claudius Régaud Institute, Toulouse, France.

[4]Hepatogastroenterology and digestive oncology department, Hôpital Européen Georges Pompidou, Paris, France.

Corresponding author:

Momar Diouf

Direction de la Recherche Clinique et de l'Innovation. Centre Hospitalier Universitaire Amiens Nord, 1, Place Victor Pauchet, F-80054 Amiens cedex, France.

E-mail: diouf.momar@chu-amiens.fr

Tel: +33 3 22 668 139

Fax: +33 3 22 667 911

Word count: 2529

Number of tables: 2

Number of figures: 4

Abstract

Background:

Eastern Cooperative Oncology Group Performance Status (ECOG PS) is currently an important parameter in the choice of treatment strategy for metastatic pancreatic adenocarcinoma (mPA) patients. However, previous research has shown that patients' self-reported quality of life provides additional prognostic information in homogeneous groups of patients with respect to ECOG PS.

The aim of this study was to identify quality of life scales with independent prognostic value in mPA and to propose prognostic groups for these patients.

Methods:

We analysed data from ninety-eight chemotherapy-naive patients with histologically proven mPA recruited from 2007 to 2011 in the FIRGEM phase II study. The FIRGEM study was designed to demonstrate the superiority of sequential treatment with FOLFIRI.3/Gemcitabine over Gemcitabine alone.

Quality of life data were assessed with EORTC QLQ-C30.

A random survival forest methodology was used to impute missing data and to identify important major prognostic factors for overall survival.

Results:

Baseline quality of life assessment was completed by 60% of patients (59/98) using the EORTC QLQ-C30.

Fourteen prognostic variables were identified. The three most important prognostic variables were fatigue, role functioning and appetite loss, followed by ASAT and CA19-9.

The model's discriminative power assessed by Harrell's C-statistic was 0.65.

Conclusions:

Quality of life scores have a prognostic value in mPA patients with good performance status (ECOG PS of 0/1). Moreover, the patient's fatigue, self-perception of daily activities and appetite loss were more reliable prognostic indicators than clinical and laboratory variables.

These quality of life scores, especially the fatigue score, should be urgently included for prognostic assessment of mPA patients (with good ECOG PS).

Background:

Pancreatic adenocarcinoma is the 13th most common cancer worldwide with a growing incidence in Europe [1]. More than 80%[2] of patients have non-curable disease at diagnosis and most of them have distant metastasis; this high percentage of patients with metastatic pancreatic adenocarcinoma (mPA) at diagnosis justifies the search for an accurate staging system for this subpopulation.

As in other types of cancer, accurate staging of mPA is essential to evaluate the acceptable degree of aggressiveness of treatment for a given patient. However, the existing TNM classification [3] for pancreatic adenocarcinoma concerns all stages and does not allow sub-classification for mPA patients for whom no classification system exist. Such a staging system would help personalization of therapy. Presently, there exist no classification systems for patients with metastatic pancreatic cancer.

Currently, in patients with metastatic disease, the National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology (ESMO) [4] recommend treatment based on the patient's Eastern Cooperative Oncology Group Performance Status (ECOG PS). The patient's well-being is therefore a decisive factor in the choice of treatment.

However, several studies [5][6][7][8] have shown that health-related quality of life (QoL) also constitutes a prognostic factor in homogeneous groups of patients with respect to ECOG PS, suggesting that QoL scores might capture complementary prognostic information not contained in the ECOG PS reported by the clinician.

Besides establishing the prognostic value of QoL scores, their incorporation into existing or new classification systems could help clinicians to take account of the patient's perceived health.

Standard statistical analysis for prognostic factors in time to event studies often use Cox proportional hazard model. However, in case of little to moderate sample size with several potential prognostic factors, Cox analysis may provide biased results.

Using data from a randomized trial for patients with mPA with a good performance status (ECOG PS 0/1), we sought to (1) illustrate the interest of random survival forest to identify important prognostic factors including QoL cores and from this (2) to define risk groups.

Patients and methods:

Patients

Data from ninety eight ata from ninety-eight chemotherapy-naive patients with histologically proven mPA recruited from 2007 to 2011 in the FIRGEM phase II study were analysed. The FIRGEM study was designed to assess the efficacy of a sequential combination of FOLFIRI.3/Gemcitabine over Gemcitabine alone. To be included, a patient had to have well-controlled pain, neutrophils ≥ 1,500/mm^3, platelets ≥ 100,000/mm^3, Hb ≥ 9 g/dl, ASAT and ALAT < 5 ULN, serum bilirubin ≤ 1.5 ULN, normal renal function, and no known brain or bone metastases.

Data collected

Clinical and laboratory variables were recorded within seven days prior to randomization. Quality of life data was assessed within 3 weeks prior to randomization using the European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 [9], a 30-item questionnaire with 7-point Likert scale for the two items regarding global health and a 4-point Likert scale for the other 28 items. The 30 items are divided into 15 scales: a global health scale, 5 multi-item functioning scales (physical emotional, social, cognitive and role functioning), 3 multi-item symptom scales (nausea and vomiting, pain, fatigue) and 6 single-item symptom scales (diarrhoea, insomnia, dyspnoea, appetite loss, constipation and financial difficulties). The response to each scale is converted into a score ranging from 0 to 100 using a linear transformation described in the EORTC scoring manual. For global health and the functioning scales, a high score corresponds to good quality of life, whereas the opposite is true for the symptom components.

Statistical methods:

Patients with available quality of life data were used in the standard Cox analysis while all randomized patients were used in the random forest analysis and in the Cox model with multiple imputations.

Baseline characteristics were expressed as median and range for continuous variable. Qualitative variables were summarized using count and percentages. Patients characteristics were compared to check whether bias selection occurred according to the availability of QoL data at baseline (Kruskall-wallis test for continuous data and fisher exact test or chi—2 for categorical variable).

Overall survival was defined as the time from randomization to death from any cause or last follow-up Survival rates were estimated using Kaplan-meier method and describe using median with 95% CI.

We adopted a random survival forest (RSF) methodology to identify prognostic factors for overall survival (OS).

In some situations like a high frequency of noise variables, RSF is known to outperform the traditional stepwise Cox model [10].

RSF analysis was based on the methodology described by Ishwaran et al.[10] using 15 clinical and laboratory variables as well as the 11 QoL components (global quality of life was not selected according to the recommendations of Van Steel et al.[11]). A total of 26 candidate predictors with univariate Cox p value<0.50 were therefore considered. Briefly, a RSF was constructed with 2000 trees; each tree was developed using a bootstrap sample of the original data with a random subset of 10 candidate predictors selected for splitting at each node (instead of the default square root of the 27 candidate variables), as, in the presence of highly correlated variables such as QoL scores, Strobl et al.[13] recommended increasing the number of random variables used for splitting a node to avoid spurious results. The variable maximizing the log-rank statistic using 10 randomly chosen splitting points was chosen as splitting variable. The splitting process continued for as long as the terminal nodes had fewer than 3 events. The out-of-bag (OOB) cumulative hazard function (CHF) for each patient was the average of the cumulative hazard function (CHF) across the nodes containing that patient for trees constructed with data excluding that patient.

Prognostic factors for overall survival were studied by means of the minimal depth rule [14]. The minimal depth (min-depth) for a splitting variable evaluates the minimal distance between that variable and the root node. The lower the min-depth, the more informative is the variable.

A variable with an average minimal depth (across the 2000 trees in the forest) less than the average minimal depth under the null hypothesis of no effect were declared to provide prognostic information [14].

In addition to variable selection and ranking according to their depth, we calculated variable importance (VIMP), which measures the increase (or decrease) in prediction error when a variable is “noised-up”. A variable is noised-up by random permuting its values[12].

Harrell’s C-index was calculated to assess the discriminative power of the model using OOB data [10].

Determination of risk groups was based on terciles of the ensemble mortality for each patient and the three Kaplan-Meier curves were plotted. The ensemble mortality

assesses the cumulative risk of death and is defined as the sum of the CHF across the different time points Imputation of missing data was based on the method described by Ishwaran et al. [10].

To facilitate interpretation and prediction, the important prognostic factors identified by the random forest were then used to build a conditional inference tree[14] (with the ctree function in the party package), which is known to be less biased than standard unconditional trees (cart function in the rpart package).

For purpose of comparison between RSF analysis and standard modelling, we also performed a multivariate Cox analysis with backward elimination including all variables with univariate p value<10% in the population with complete QoL scores and in the whole population after multiple imputations with five replications. Hazard ratio (HR) and the corresponding 95% two-sided confidence intervals (95%CI) were computed as well as Schemper statistic and Harrell C-index.

P-values were two-sided and variables with $p<0.05$ were considered significantly associated with OS in multivariate analysis.

Statistical analyses were performed with SAS® software (version 9.2, SAS Institute Inc., Cary, NC) and R.3.1.0 software (free) using the randomForestSRC package (http://cran.r-project.org/web/packages/randomForestSRC/index.html).

List of variables considered in the RSF analysis:

Clinical and laboratory variables:

Age, tumor size, number of metastatic sites, liver metastasis, lung metastasis, Eastern Cooperative Oncology Group Performance Status (ECOG PS), carbohydrate antigen 19-9(CA19-9), lactate dehydrogenase (LDH), aspartate amino transferase (ASAT), alanine amino transferase (ALAT), neutrophils, Alkaline phosphatase, hemoglobin, leucocytes.

Quality of life variables:

Role, cognitive, social and physical functioning, followed by fatigue, dyspnea, insomnia, nausea/vomiting, pain and appetite loss.

Results:

Description of the population*:*

Between 2007 and 2011, 98 patients with a mean age of 62 years were enrolled. The majority were men (60%: 59/98) most patients had an ECOG PS 1 (68%: 67/98) and liver metastasis was frequent (75%:74/98).

The median [range] values for CA19-9, LDH, alkaline phosphatase and ASAT were 920 IU/L [0.6; 913750], 291 IU/L [96; 5022], 151 IU/L [42; 1811] and 35 IU/L [8; 187], respectively.

Baseline QoL scores were recorded for 59 patients (60%) and are summarized in Table 1.

The median pain score was 33.33, while 25% (14/55) of patients had a pain score > 66.66.

About one half of patients had global quality of life, physical functioning, role functioning, fatigue and appetite loss scores greater than 41.7, 86.7, 83.3, 44.4 and 33.3, respectively.

After a median follow-up of 23 months, eighty one patients had died (83%). The median overall survival was 8.9 [6.4; 10.03]. There was no correlation between the availability of QoL data and overall survival (HR=0.73 (95%CI=[0.47; 1.14]) available vs. not available, p=0.1690)) Median Overall survival were 9.6 months (95%CI=[8.4; 13.0]) and 7.1 months (95%CI=[4.8; 10.8]) for patients with and without QoL data, respectively.

Variable	*Mean*	*Standard deviation*	*Minimum*	*Maximum*	*N*	*1st quartile*	*median*	*3th quartile*
Age (years)	62.30	8.39	38.81	76.34	98	57.81	62.72	69.01
IMC (kg/m^2)	24.34	5.07	15.99	47.87	98	21.39	23.46	26.26
ASAT (IU/L)	44.61	35.42	8.00	187.00	95	21.00	35.00	52.00
Leucocytes (/mm$^{3)}$	8793.62	4382.72	85.00	36500.00	98	6400.00	7955.00	10200.00
Neutrophils (/mm$^{3)}$	6253.25	3996.43	1800.00	32850.00	97	4090.00	5445.00	6930.00
Serum creatinine (µmol/L)	71.67	16.04	39.00	108.00	96	60.00	70.00	84.00
ALAT (IU/L)	61.81	60.94	8.00	348.00	95	24.00	45.00	71.00
Protombin (%)	87.64	16.20	19.00	122.00	80	80.00	88.50	100.00
Platelets (/mm^3)	279.40	107.63	94.00	634.00	97	196.00	270.00	345.00
Hemoglobin (g/dl)	12.87	1.72	7.90	16.50	97	11.80	12.90	14.10
Bilirubin (µmol/L)	18.18	27.73	1.00	227.00	93	7.40	12.00	20.00
Alcalines phosphatises (IU/L)	277.92	333.16	42.00	1811.00	95	89.00	151.00	301.00
Glycemia (mmol/L)	6.74	2.60	0.70	15.00	54	5.37	6.00	7.55
CA199 (IU/L)	26031.62	112173.62	0.60	913750.00	91	98.00	920.00	6309.40
LDH (IU/L)	435.89	720.50	96.00	5022.00	45	190.00	291.00	452.00
Sum of tumor diameters (cm)	84.83	45.94	15.00	260.00	70	51.00	78.50	107.00
Global health	41.97	22.74	0.00	100.00	55	25.00	41.67	50.00
Role functioning	72.03	31.32	0.00	100.00	59	50.00	83.33	100.00
Physical functioning	77.65	22.32	26.67	100.00	54	66.67	86.67	93.33
Cognitive functioning	81.25	24.43	0.00	100.00	56	66.67	83.33	100.00
Social functioning	73.81	28.93	0.00	100.00	56	50.00	83.33	100.00
Emotional functioning	64.26	22.10	8.33	100.00	52	50.00	66.67	83.33
Pain	38.18	28.99	0.00	100.00	55	16.67	33.33	66.67
Fatigue	47.47	31.02	0.00	100.00	55	33.33	44.44	66.67
Appetite loss	49.15	40.76	0.00	100.00	59	0.00	33.33	100.00
Nausea and vomiting	13.45	19.01	0.00	66.67	57	0.00	0.00	33.33
Dyspnoea	15.25	27.21	0.00	100.00	59	0.00	0.00	33.33
Insomnia	37.85	35.26	0.00	100.00	59	0.00	33.33	66.67
Constipation	35.15	39.24	0.00	100.00	55	0.00	33.33	66.67
Diarrhoea	10.71	24.71	0.00	100.00	56	0.00	0.00	0.00
Financial difficulties	9.94	23.54	0.00	100.00	57	0.00	0.00	0.00

Table 1: Demographical, clinico-biological and quality of life data for patients randomized in the FIRGEM study.

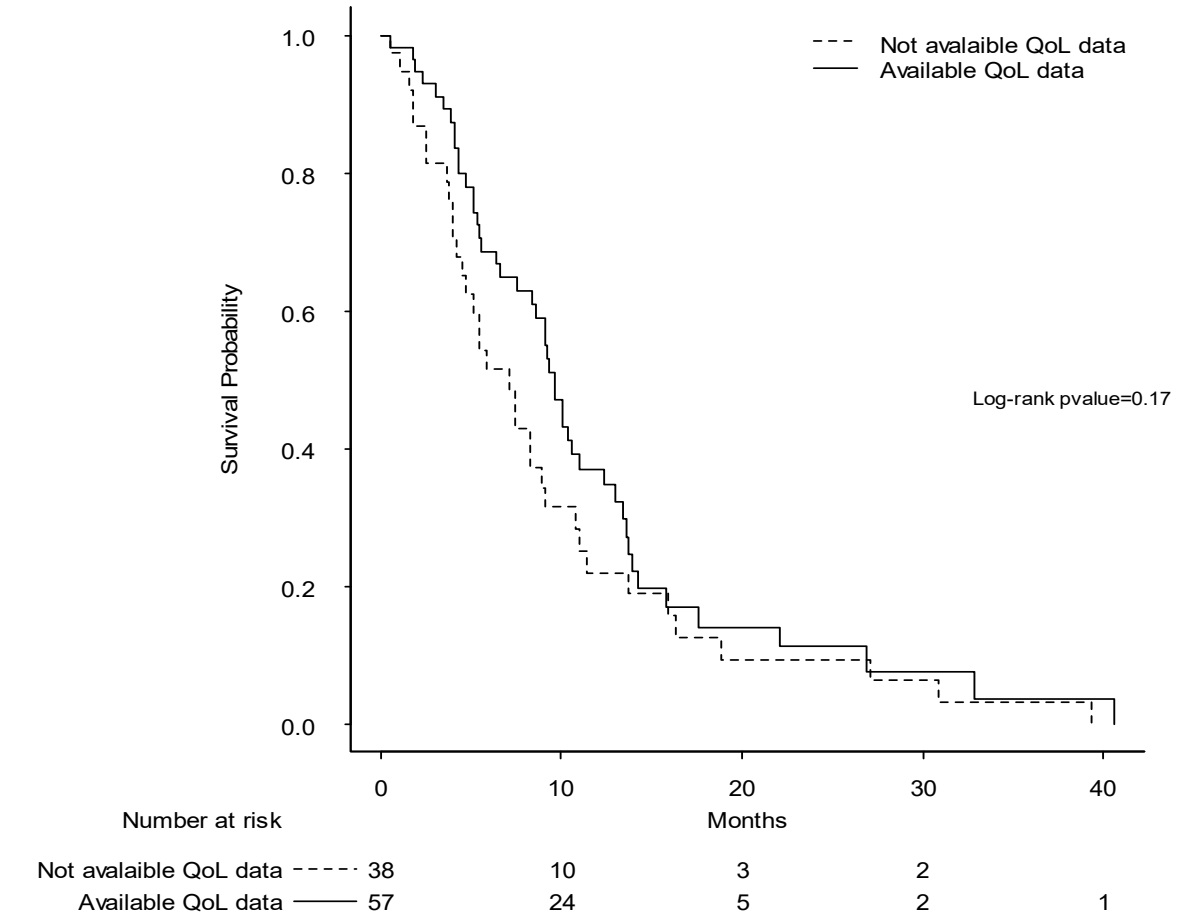


Figure 1: Overall survival according to QoL availability**.** Median survival was 9.6 [8.4 – 13.0] and 7.1 [4.8 – 10.8] for patients with and without QoL data, respectively.

RSF prognostic analysis:

Results of RSF analysis are summarized in Table 2. ASAT, CA19-9, LDH, neutrophils, alkaline phosphatase, haemoglobin, tumour size and leukocytes provided prognostic information.

Quality of life scales with prognostic value were fatigue, role functioning, appetite loss, physical functioning, cognitive functioning and insomnia.

The three prognostic groups are shown in Figure 2 and Table 2 compares the important prognostic variables between these three risk groups.

Median survival was approximately 4, 9 and 18 months for the poor, intermediate and good prognosis groups, respectively.

The tree constructed with important prognostic factors resulted in one single split (a depth of one) for fatigue score with a splitting point of 46.98 Figure 3).

Figure 4A and 4B show the two prognostic groups according to the median ensemble mortality and fatigue score, respectively.

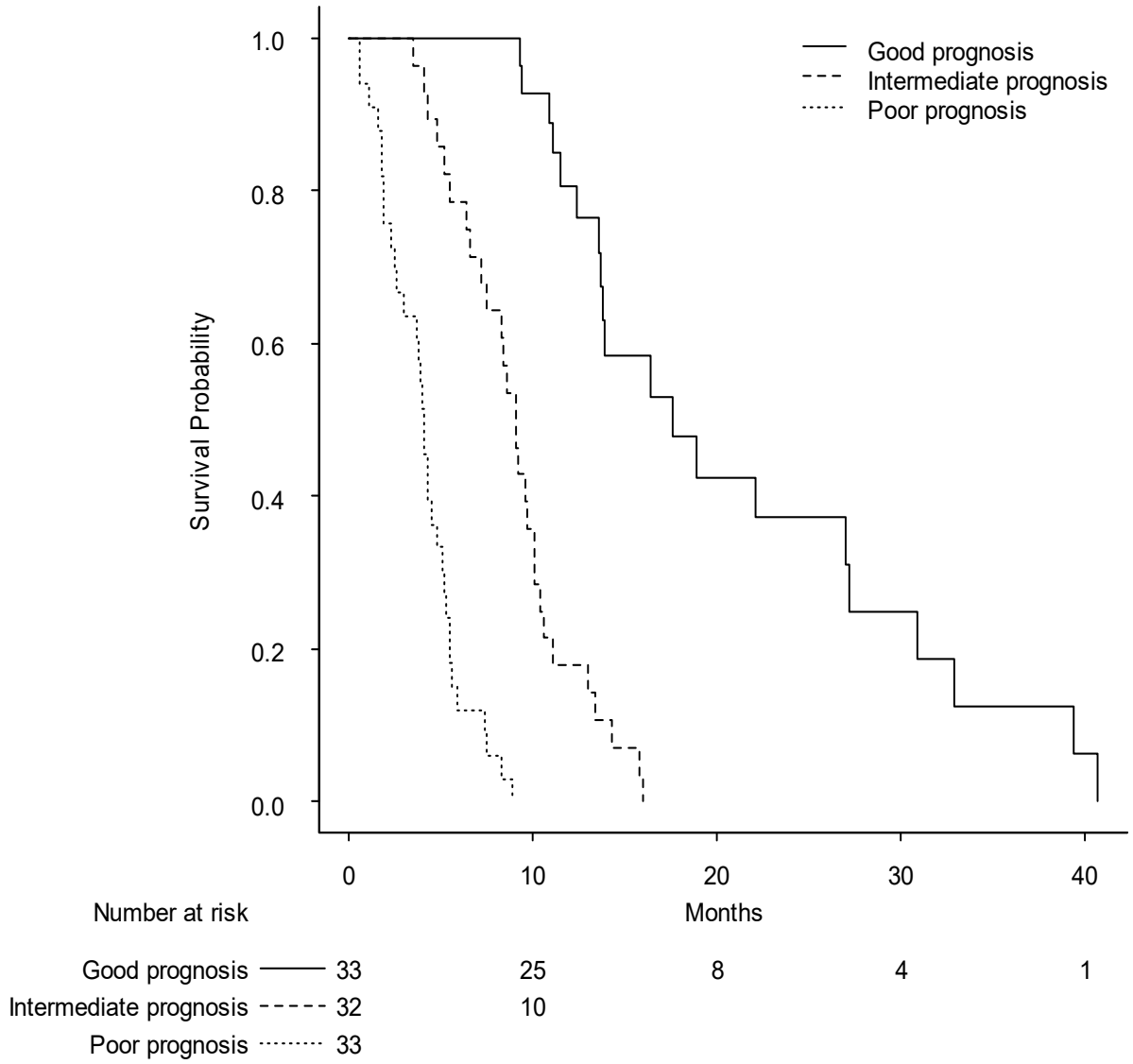


Figure 2: Overall survival (months) for good, intermediate and poor prognosis according to the tertiles of the ensemble mortality derived from the random survival forest analysis.

Median survival was 17.6 [13.8 – 30.9], 9.1 [7.5 – 10.4] and 4.1 [3.0 – 5.2] for good (solid line), intermediate (dash line) and poor prognosis (dotted line), respectively.

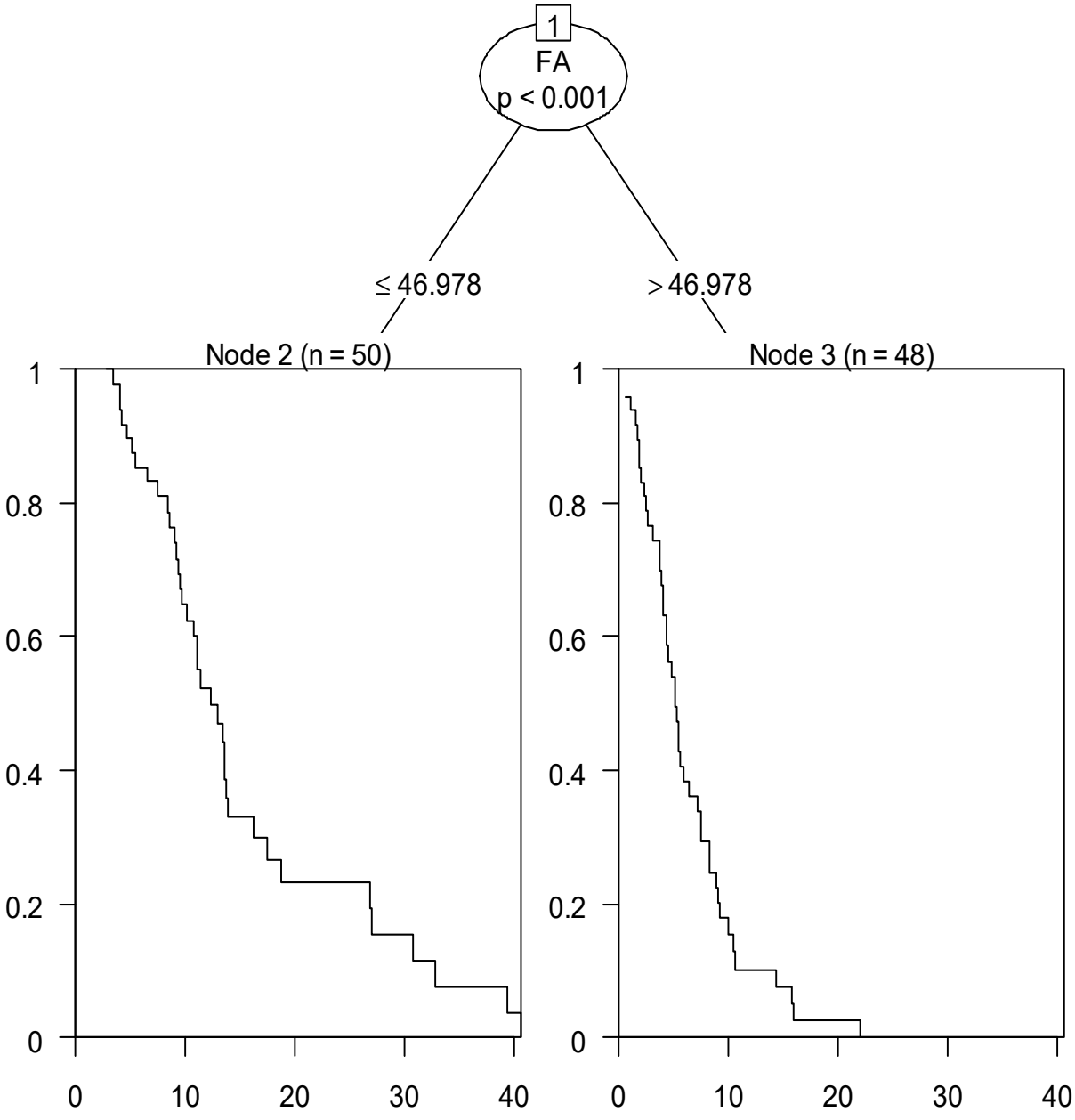
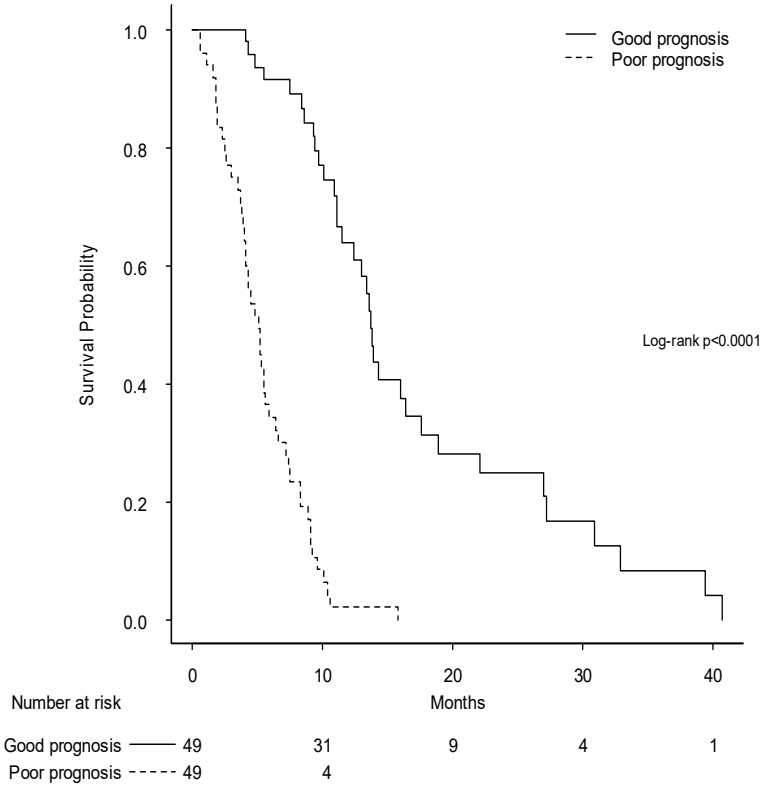


Figure 3: Results of the conditional inference tree constructed with important variables identified by the random forest analysis. The left survival curve concerns the good prognosis group and the left curve illustrates the survival distribution for the poor prognosis group with a cutoff value of 47.0 for the fatigue score.



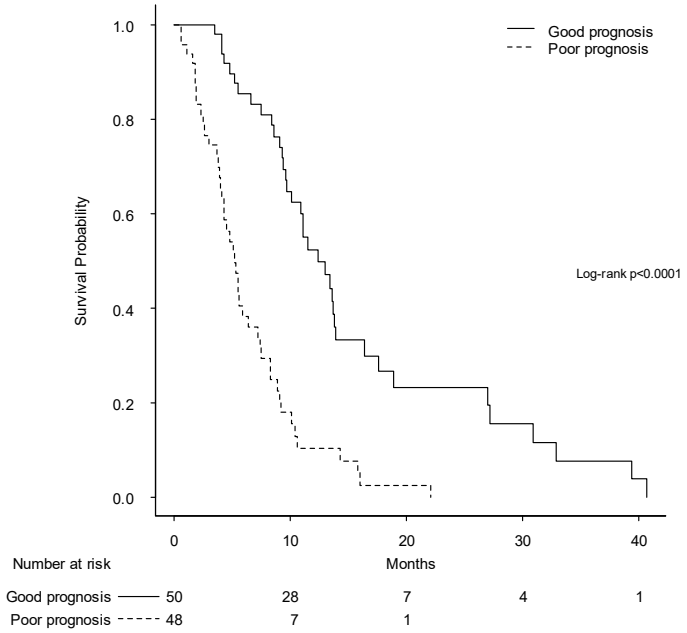


Figure 4: Overall survival (months) for good and poor prognosis according to the median ensemble mortality derived from the random survival forest analysis and from dichotomization of the fatigue score, respectively.

A) Median survival was 13.7 [12.3 – 18.8] and 5.1 [4.1 – 6.4] for good (solid line) and poor prognosis (dashed line), respectively, derived from the median ensemble mortality.

B) Median survival was 12.3 [10.1 – 16.3] and 5.2 [4.3 – 7.4] for good (solid line) and poor prognosis (dashed line), respectively, derived from the fatigue score with a cutoff value of 47.0.

Standard Cox analysis with and without imputation of missing data:

Table 3 summarizes the results of univariate and multivariate Cox models:

In univariate analysis, eight clinical and laboratory variables and nine QoL scales were significantly associated with OS at a 10% alpha level.

In multivariate complete case analysis, only role functioning and insomnia were independent prognostic factors for OS with HR=0.980 (95%CI=[0.966; 0.993]) and HR=1.021 (95%CI=[1.007; 1.036]), respectively. The Harrell C-index was 0.71 (95%CI=[0.58; 0.85]).After multiple imputations, no variable was selected among the five replications.

variables	Depth	VIMP	Group 1	Groupe2	Group 3	Pvalue
Fatigue	4.05	0.021	31.34±20.8 33.3 [0.0 ; 100.0] 33.3 [26.4 ; 39.3]	48.2±21.6 44.4 [11.1 ; 100.0] 44.4 [33.3 ; 63.1]	73.2±16.0 70.4 [44.4 ; 100.0] 70.4 [62.3; 79.2]	<0.0001
Role functioning	4.45	0.012	90.5±13.9 89.6 [33.3 ; 100.0] 89.6 [83.3; 100.0]	69.5±23.6 80.1 [0.0 ; 100] 80.1 [50.0 ; 83.3]	52.8±25.9 54.5 [0.0 ; 100.0] 54.5 [44.1 ; 66.7]	<0.0001
Appetite loss	4.53	0.011	33.2±29.6 33.3 [0.0 ; 100.0] 33.3 [0.0 ; 51.6]	48.2±35.6 40.1 [0.0 ; 100.0] 40.1 [30.6 ; 71.8]	77.8±16.00 73.7 [33.3 ; 100.0] 73.7 [67.8 ; 100.0]	<0.0001
ASAT	4.65	0.004	32.6±29.6 21.0 [10.0 ; 143.0] 21.0 [17.0 ; 38.0]	45.1±21.4 42.4 [13.0 ; 108.0] 42.4 [28.5 ; 53.5]	56.5±45.5 46.0 [8.0 ; 187.0] 46.0 [23.0 ; 66.0]	0.0013
CA19-9 (UI/L)	4.76	0.005	4119±9670 480[0.6 ; 44775] 480 [98 ; 3738]	7373±19440 257[2.5 ; 100663] 257 [64.5 ; 4216]	71375±179564 8233[1; 913750] 8233[1510; 55552]	<0.0001
Physical functioning	4.79	0.013	88.0±10.4 90.1 [53.3 ; 100.0] 90.1 [85.4 ; 93.3]	78.2±15.3 80.7 [40.0 ; 100.0] 80.7 [71.9 ; 87.9]	61.4±18.2 62.5 [26.7 ; 100.0] 62.5 [54.4 ; 67.0]	<0.0001
LDH (IU/L)	4.81	0.006	315.2±115.1 302.0 [142.0 ; 682.6] 302.0 [247.0 ; 360.0]	400.0±160.6 365.4 [133.0 ; 766.0] 365.4 [316.6 ; 512.3]	654.3±828.0 512.9 [96.0 ; 5022.0] 512.9 [300.0 ; 716.0]	0.0014
Neutrophils(/mm^3)	4.81	0.007	5005±3001 4230[2300; 15152] 4230[3300 ; 5092]	6219±2280 6045[1800 ; 10400] 6045[4410 ; 8070]	7528±5531 6020[3160 ; 32850] 6020[5000 ; 7300]	0.0002
Alcalines Phosphatases (IU/L)	4.85	0.003	175±192 100[42; 1069] 100 [85 ; 226]	282±330 183 [59; 1811] 183 [101; 333]	387±403 236 [63; 1758] 246 [89; 528]	0.0233
Hemoglobin (g/dl)	4.88	0.002	13.1±1.4 13.2 [10.1 ; 16.5] 13.2 [12.0 ; 14.2]	12.9±1.8 12.8 [9.4 ; 16.4] 12.8 [11.7 ; 14.3]	12.6±1.8 12.8 [7.9 ; 15.9] 12.8 [11.6 ; 14.0]	0.5922
Cognitive functioning	4.95	0.007	90.0±11.7 89.2 [50.0 ; 100.0] 89.2 [83.3 ; 100.0]	82.9±20.3 84.1 [40.0 ; 100.0] 84.1 [80.7 ; 100.0]	65.1±18.7 66.7 [0.0 ; 100.0] 66.7 [56.7 ; 75.5]	<0.0001
Insomnia	5.06	0.004	27.4±26.9 32.4 [0.0 ; 100.0] 32.4 [0.0 ; 37.3]	40.9±28.6 33.3 [0.0 ; 100.0] 33.3 [33.1 ; 66.7]	47.4±24.3 42.9 [0.0 ; 100.0] 42.9 [33.3 ; 50.0]	0.0023
Tumor size (cm)	5.11	0.003	70.9±30.6 72.6 [21.0 ; 152.0] 72.6 [47.0 ; 87.1]	89.6±31.3 90.0 [15 ; 157.0] 90.0 [75.5 ; 103.5]	99.6±49.0 94.2 [24.0;260.0] 94.0 [62.0 ; 110.0]	0.0100
Leucocytes (/mm^3)	5.15	0.003	7576±3509 6630 [4100; 20900] 6630 [6000 ; 7600]	8612±3075 8600 [85 ;14400] 8600 [6690 ; 10600]	10188.3±5760 8770 [5030; 36500] 8770 [7600 ; 10300]	0.0012

Table 2: Results of the random survival forest analysis and comparison of the three prognostic groups built using the ensemble mortality's tertiles.

	Univariate analysis		Multivariate analysis	
variables	HR (95%CI)	pvalue	HR (95%CI)	pvalue
Age	1.01 [0.99; 1.04]	0.2354		
Sex	0.98 [0.62; 1.54]	0.9160		
BMI	0.996 [0.954; 1.038]	0.8360		
ECOG PS	1.61 [0.99; 2.61]	0.0548		
Tumor size	1.006 [1.000; 1.012]	0.0447		
Localization				
Tail vs. head	0.99 [0.55; 1.78]			
Body vs. head	1.47 [0.83; 2.61]	0.5773		
Number of metastatic site	1.26 [0.92; 1.71]	0.1459		
Liver metastasis	1.49 [0.89; 2.51]	0.1307		
Lung metastasis	1.29 [0.76; 2.17]	0.3421		
Peritonis metastasis	0.91 [0.54; 1.53]	0.7330		
Node involvement	0.96 [0.49; 1.88]	0.9102		
bilirubin	1.000 [0.993; 1.007]	0.9639		
LDH	1.000 [1.000; 1.001]	0.4698		
CA19-9	1.000 [1.000; 1.000]	0.1009		
neutophyles	1.000 [1.000; 1.000]	0.0046		
glycemy	1.008 [0.901; 1.127]	0.8923		
ASAT	1.009 [1.003; 1.015]	0.0038		
ALAT	1.004 [1.000; 1.008]	0.0320		
Prothrombin	0.999 [0.982; 1.017]	0.9430		
Alkaline phosphatase	1.001 [1.000 ; 1.001]	0.0238		
Hemoglobin	0.915 [0.792; 1.056]	0.2244		
Serum creatinine	0.997 [0.982; 1.013]	0.7411		
Leucocytes	1.000 [1.000 ; 1.000]	0.0234		
Global helath	1.000 [0.996; 1.027]	0.1594		
Role functioning	0.984 [0.974; 0.994]	0.0019	0.980 [0.966; 0.993]	0.0029
Physical functioning	0.983 [0.969; 0.997]	0.0174		
Emotional functioning	0.987 [0.984; 1.010]	0.6676		
Cognitive functioning	0.988 [0.977; 1.000]	0.0547		
Social functioning	0.991 [0.982; 1.001]	0.0785		
Fatigue	1.015 [1.005; 1.025]	0.0027		
Diarrheoa	1.001 [0.991; 1.012]	0.7857		
Dynpnoea	1.004 [0.995; 1.014]	0.3876		
Finacial difficulties	1.003 [0.991; 1.016]	0.6127		
Insomnia	1.012 [1.003; 1.021]	0.0108	1.021 [1.007; 1.036]	0.0028
Nausea/Vomiting	1.007 [0.991; 1.023]	0.3925		
Pain	1.011 [1.001; 1.021]	0.0359		
Appetite loss	1.012 [1.004; 1.020]	0.0023		
Constipation	1.007 [0.999; 1.015]	0.0769		

In multivariate MI analysis with 5 replications, none of the variables was selected by backward elimination method.

Table 3: Results of multivariate and univariate Cox analyses.

Discussion:

By increasing order of importance, fatigue, role functioning, appetite loss, ASAT, CA19-9, physical functioning, LDH, neutrophils, alkaline phosphatase, haemoglobin, cognitive functioning, insomnia, tumour size and leukocytes were prognostic factors for overall survival according to the min-depth rule. In contrast, according to VIMP, the corresponding rank was: fatigue, physical functioning, role functioning, appetite loss, neutrophils, cognitive functioning, LDH, CA19-9, ASAT, insomnia, alkaline phosphatase, tumour size, leukocytes and haemoglobin. All these variables met the proposed threshold value of 0.002 for VIMP [10] which measures the relative importance of each variable for prognosis assessment.

With respect to the observed difference between the three prognostic groups, the variable ranking using VIMP was more coherent with the p value than the min-depth ranking. We believe that, in this particular case, the VIMP rule is more efficient than the min-depth rule for selection of important predictors.

More importantly, the top three variables were QoL scores (including fatigue and role functioning) regardless of the variable selection method (min-depth or VIMP). The tree constructed using conditional inference[14] retained fatigue as the single important variable for prognostic assessment with a score >47 associated with a very poor prognosis (median survival of 5 months). This means that, although the other variables had a prognostic value, the patient's fatigue score contained almost all of the prognostic information.

This result is confirmed by the similarity of the two prognostic groups defined by fatigue and the ensemble mortality (derived from the RFS analysis): 88% (43/49) of patients in the poor prognosis group defined by the ensemble mortality had a fatigue score >47 and 90%(44/49) of patients in the good prognosis group defined by the ensemble mortality had a fatigue score ≤47. The fatigue score therefore had an excellent sensitivity to detect either patients with poor or good prognosis.

Role functioning and insomnia were the only variables selected in the multivariate Cox analysis using complete case data (patients with available QoL scores). On the other hand, after multiple imputations with five replications, no variable was selected among the five multivariate Cox models constructed with each of the five new databases. Moreover, CA19-9, which was one of the most important factors according to random forest analysis, was not selected even by univariate Cox analysis at a 10% alpha level. These results confirm the well-known test/estimation[15] problem of stepwise model selection when assessing independent predictors. While the main goal of prognostic assessment is an estimation problem, the stepwise method performs a test before estimation leading to elimination of important prognostic factors from the final model and overestimation of the effect of the variables retained. RSF analysis avoids these pitfalls of stepwise methods and may be an attractive alternative, especially when only a small sample is available.

LDH and CA19-9 have already been identified as important predictors for OS by Haas et al. [16] and the prognostic value of ASAT and alkaline phosphatase in our cohort might be related to the high percentage of patients with liver metastases (75%).

ASAT and alkaline phosphatase confirmed the prognostic value of liver injury, as shown by Haas et al.[16] who reported the prognostic value of serum bilirubin. The lack of prognostic importance of bilirubin could possibly be related to our homogeneous population with respect to bilirubin (only patient with bilirubin ≤1.5 ULN was included in this study).

Bernard el al. [6] found that tiredness (fatigue) was an independent prognostic factor for OS in metastatic pancreatic cancer, while Gourgou et al.[7] found that physical functioning was a predictor for OS. Our results confirmed these two findings, although the prognostic value of fatigue appeared to be greater than that of physical functioning.

Previous research [7][6] has demonstrated the prognostic value of pain in patients with advanced pancreatic cancer. Given that only patients with manageable pain were included, the non-significant prognostic value of QLQ-C30 pain score is therefore logical.

The marked importance of QoL data for prognosis in our cohort confirms the hypotheses proposed by Gotay [17] and Mauer [18] that QoL scores may detect disease progression earlier than conventional clinical and laboratory factors.

The high proportion of missing QoL data is a limitation to our study, although the imputation algorithm of Ishwaran et al.[10] appeared to be reliable.

The restricted inclusion and exclusion criteria limit the population concerned by our findings.

Despite the known accuracy of random forest methodology for predictors selection, our results need to be validated in an independent cohort of patients with mPA.

Conclusion:

The RSF technique was more efficient than standard Cox analysis for screening important prognostic factors in our study. RSF is thus a promising technique in prognostic studies including QoL data. Although subjective, self-reported QoL scores provided additional important information regarding the patient’s survival and outperformed clinical and laboratory factors.

Given their high prognostic value, these QoL scores, especially the fatigue score, should be considered in the pre-treatment evaluation of mPA patients in addition to the ECOG PS. The cutoff value of 47 for the fatigue score should also be considered

for stratification or inclusion/exclusion criteria in randomized clinical trials including patients with mPA.

References:

1. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. CA Cancer J Clin 55: 74–108.

2. Hidalgo M (2010) Pancreatic cancer. N Engl J Med 362: 1605–1617. doi:10.1056/NEJMra0901557.

3. International Union against Cancer (2010) TNM classification of malignant tumours. 7th ed. Sobin LH, Gospodarowicz MK, Wittekind C, editors Chichester, West Sussex, UK ; Hoboken, NJ: Wiley-Blackwell. 309 p.

4. Seufferlein T, Bachet JB, Van Cutsem E, Rougier P, ESMO Guidelines Working Group (2012) Pancreatic adenocarcinoma: ESMO-ESDO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol Off J Eur Soc Med Oncol ESMO 23 Suppl 7: vii33–vii40. doi:10.1093/annonc/mds224.

5. Diouf M, Filleron T, Barbare J-C, Fin L, Picard C, et al. (2013) The added value of quality of life (QoL) for prognosis of overall survival in patients with palliative hepatocellular carcinoma. J Hepatol 58: 509–521. doi:10.1016/j.jhep.2012.11.019.

6. Bernhard J, Dietrich D, Glimelius B, Hess V, Bodoky G, et al. (2010) Estimating prognosis and palliation based on tumour marker CA 19-9 and quality of life indicators in patients with advanced pancreatic cancer receiving chemotherapy. Br J Cancer 103: 1318–1324. doi:10.1038/sj.bjc.6605929.

7. Gourgou-Bourgade S, Bascoul-Mollevi C, Desseigne F, Ychou M, Bouché O, et al. (2013) Impact of FOLFIRINOX compared with gemcitabine on quality of life in patients with metastatic pancreatic cancer: results from the PRODIGE 4/ACCORD 11 randomized trial. J Clin Oncol Off J Am Soc Clin Oncol 31: 23–29. doi:10.1200/JCO.2012.44.4869.

8. Braun DP, Gupta D, Staren ED (2013) Longitudinal health-related quality of life assessment implications for prognosis in stage IV pancreatic cancer. Pancreas 42: 254–259. doi:10.1097/MPA.0b013e31825b9f56.

9. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, et al. (1993) 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 85: 365–376.

10. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS (2008) Random survival forests. Ann Appl Stat 2: 841–860. doi:10.1214/08-AOAS169.

11. Van Steen K, Curran D, Kramer J, Molenberghs G, Van Vreckem A, et al. (2002) Multicollinearity in prognostic factor analyses using the EORTC QLQ-C30: identification and impact on model selection. Stat Med 21: 3865–3884. doi:10.1002/sim.1358.

12. Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. BMC Bioinformatics 9: 307. doi:10.1186/1471-2105-9-307.

13. Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS (2010) High-Dimensional Variable Selection for Survival Data. J Am Stat Assoc 105: 205–217. doi:10.1198/jasa.2009.tm08622.

14. Hothorn T, Hornik K, Zeileis A (2006) Unbiased Recursive Partitioning: A Conditional Inference Framework. J Comput Graph Stat 15: 651–674. doi:10.1198/106186006X133933.

15. Steyerberg EW (2009) Clinical prediction models: a practical approach to development, validation, and updating. New York, NY: Springer. 497 p.

16. Haas M, Heinemann V, Kullmann F, Laubender RP, Klose C, et al. (2013) Prognostic value of CA 19-9, CEA, CRP, LDH and bilirubin levels in locally advanced and metastatic pancreatic cancer: results from a multicenter, pooled analysis of patients receiving palliative chemotherapy. J Cancer Res Clin Oncol 139: 681–689. doi:10.1007/s00432-012-1371-3.

17. Gotay CC, Kawamoto CT, Bottomley A, Efficace F (2008) The prognostic significance of patient-reported outcomes in cancer clinical trials. J Clin Oncol Off J Am Soc Clin Oncol 26: 1355–1363. doi:10.1200/JCO.2007.13.3439.

18. Mauer M, Bottomley A, Coens C, Gotay C (2008) Prognostic factor analysis of health-related quality of life data in cancer: a statistical methodological evaluation. Expert Rev Pharmacoecon Outcomes Res 8: 179–196. doi:10.1586/14737167.8.2.179.

V. Discussion :

Dans ce travail, la valeur pronostique de la qdv a été établie dans trois localisations cancéreuses en situation avancée : le cancer colorectal, le CHC et l'ACP.

Concernant le cancer colorectal, la mobilité et la douleur mesurées par le questionnaire EUROQOL EQ-5D sont des facteurs pronostiques indépendants de survie permettant d'améliorer de façon significative les performances des systèmes de classifications de Köhne et du GERCOR. La valeur pronostique de la mobilité est cohérente avec les résultats de Maisey et al.[42] (dimensions de bien-être physique et d'activité quotidienne), Lis et al.[48] (bien-être physique) et Efficace et al.[41] (bien-être social) puisqu'une mobilité réduite impacte directement les performances physiques, l'activité quotidienne et la vie sociale. Au contraire, seule l'étude de Maisey n'ayant pas mis tous les scores de qdv en compétition dans un même modèle multivarié a retrouvé la douleur comme facteur pronostique de survie globale.

Pour ce qui est du CHC, seule l'activité quotidienne est un facteur indépendant de survie en analyse multivariée alors que la fatigue et la diarrhée évaluées avec le QLQ-C30 sont les scores de qdv qui contribuent le plus à l'amélioration des performances des systèmes de classification pronostique habituels (BCLC, CLIP, GRETCH et BoBar). L'absence du score d'activité quotidienne pour l'amélioration de ces systèmes de classification pourrait s'expliquer par la présence de l'IP-OMS dans les systèmes BCLC et BoBar et du score de Karnosfsky dans le système de classification GRETCH, mais aussi par un fort impact de la fatigue sur l'activité quotidienne (Braun et al. [40] ont trouvé une corrélation de -0.80 entre ces deux dimensions de la qdv); ce sont donc les scores de symptômes qui améliorent le plus les systèmes de classification étudiées comparés aux dimensions fonctionnelles de la qdv. Le score d'activité quotidienne a déjà été retrouvé comme facteur pronostique indépendant de survie par Yeo et al. [45] en plus de la perte d'appétit. La forte corrélation entre la fonction physique et l'activité quotidienne (rho de Pearson=0.7) pourrait expliquer qu'une seule variable soit retenue dans notre modèle final, l'information pronostique étant pour une bonne partie redondante. Dans ce travail, la perte d'appétit est retrouvée comme facteur pronostique en analyse univariée mais pas en multivariée. Cela pourrait s'expliquer par la présence de la variable « albumine » (qui est corrélée à la dénutrition) dans notre module multivarié alors que cette variable n'est pas sélectionnée dans le modèle final de Yeo et al.[45] L'activité quotidienne étant une composante de l'index de Spitzer utilisé par

Bonnetain et al.[46], la valeur pronostique de cet index est concordante avec nos résultats.

Dans le cas du CHC ce travail a permis de montrer que les valeurs seuil 66.66 et 33.33 étaient optimales (en termes de séparation des patients en deux groupes pronostiques homogènes) pour les scores de fatigue et de diarrhée respectivement. En utilisant ces valeurs seuil, une mise à jour des quatre systèmes de classification pronostique ci-dessus intégrant les scores de qdv a été proposée. Cette mise à jour a permis d'améliorer la performance des systèmes de classification, plus particulièrement celle du BCLC qui est la plus utilisée même si sa performance est faible pour les patients atteints de CHC incurable. La mise en évidence dans ce travail d'un groupe de bon pronostic (en l'absence de traitement) avec 11 mois de médiane de survie équivalente à celle des patients recevant du sorafenib est cohérente avec les résultats de Yau et al. [76] et pose la question d'un éventuel réajustement de l'algorithme de traitement pour ces patients de la classe « CHC avancé ». En effet, une chimio-embolisation pourrait peut-être améliorer la survie de ces patients de bon pronostic comme cela a été montré dans la récente étude de Yau et al. [76] qui a permis de construire le 2ème système de classification associé à un algorithme de recommandation de traitement dans le CHC. L'étude de Yau et al.[76] a montré que parmi les patients de la classe BCLC-C (qui devraient recevoir du sorafenib selon l'algorithme du BCLC), un sous-groupe de patients pourrait gagner en survie s'il était traité par chimio-embolisation (la survie à 5 ans passerait de 1.7% sous sorafenib à 7.1% avec la chimio-embolisation). De même un décalage de la stratégie thérapeutique dans la classe BCLC-B pourrait améliorer la survie à 5 ans de 0% à 48.6% (chimio-embolisation versus les thérapies curatives que sont la résection, la transplantation et la destruction percutanée par radiofréquence). Cependant, un tel gain de survie devrait être prouvé au préalable dans la population européenne puisque l'étude de Yau et al.[76] inclut des patients avec CHC sur cirrhose d'étiologie majoritairement virale B et ces patients ont souvent une meilleure fonction hépatique que les patients CHC sur cirrhose d'origine alcoolique majoritaires dans notre étude. Il se pourrait donc que ces patients des sous-classes du BCLC-B ou BCLC-C qui ont un bon pronostic et qui gagneraient à avoir un traitement plus radical soient aussi ceux ayant une meilleure qdv avant l'initiation du traitement. Ainsi les sous-groupes de bon pronostic identifiés par Yau et al.[76] pourraient être identiques à ceux retrouvés dans ce travail sur le CHC qui concerne majoritairement

les patients B et C du BCLC. La qdv n'étant pas mesurée dans l'étude de Yau et al.[76], il serait intéressant que les études prospectives pour sa validation prévoient une mesure de la qdv des patients avec le QLQ-C30 et son module spécifique pour le CHC (HCC-18) afin de tester notre hypothèse.

L'étude de la valeur prédictive de la qdv dans la réponse à chacun des traitements palliatifs du CHC pourrait également aider les chercheurs à mieux expliquer l'hétérogénéité des patients des groupes BCLC-B et C et aux cliniciens de mieux personnaliser leur prise en charge.

Dans l'ACP métastatique, pratiquement seul le score de fatigue évalué avec le QLQ-C30 permet d'estimer le pronostic. Dans ce dernier cas, la méthodologie utilisée a permis de trouver une valeur seuil du score de fatigue pour une classification en bon ou mauvais pronostic facilitant ainsi l'utilisation de l'auto-perception de la fatigue dans la prise de décision thérapeutique. Cette place de la fatigue dans la classification pronostique du patient (et donc du choix du traitement) est cohérente avec l'algorithme de l'EASL pour le traitement des patients atteints de CHC pour lesquelles l'IP-OMS et la fonction hépatique (évaluée avec le score Child-Pugh) sont au sommet de l'algorithme décisionnel contrairement au stade tumoral.

De façon générale, le score de fatigue a une grande pertinence dans l'évaluation pronostique du patient comme c'est le cas dans d'autres types de cancer[82][83][84][85].

La fatigue a permis d'améliorer les performances de trois des quatre systèmes pronostiques dans le CHC avancé et explique pratiquement toute la variabilité du pronostic dans l'ACP métastatique. Dans le cas du cancer colorectal, ce travail ne permet pas d'évaluer l'influence de la fatigue sur le pronostic puisque le questionnaire EQ-5D qui n'est pas spécifique du cancer ne pose pas de question précise sur la fatigue perçue par le patient. La valeur pronostique de la fatigue dans l'ACP est cohérente avec les résultats de Robinson [50], Bernhard [51] et Gourgou [52]. La concordance de nos résultats en terme d'intérêt pronostique de la fatigue et de similarité des deux groupes pronostiques avec ceux de Robinson et al.[50] (en utilisant le FACIT-F) est particulièrement intéressante malgré la différence de questionnaires utilisés. La fatigue est le principal facteur pronostique dans les deux études et les médianes de survie sont de 5.2 et 9.1 mois respectivement pour les groupes de mauvais et bon pronostic trouvés par Robinson[50] (valeur seuil de 30 pour le FACIT-T qui varie entre 0 et 52). Dans notre travail, pour une valeur seuil de

47 du score de fatigue (QLQ-C30) les médianes de survie sont respectivement de 5.2 et 12.3 mois pour les groupes de mauvais et bon pronostic. L'importance pronostique des scores de symptôme semble donc supérieure à celle des scores fonctionnels.

Etant donnée l'importance pronostique de la fatigue et sa forte prévalence dans le cancer (entre 50 et 90%), son évaluation lors du diagnostic devrait être généralisée au moins pour les patients atteints de CHC et d'ACP avancés pour une prise en charge optimale aussi bien de la fatigue que du cancer lui-même. Cela permettrait également une prise en compte du ressenti du patient dans la classification pronostique et donc dans la stratégie thérapeutique.

Les patients inclus dans les trois essais cliniques étudiés dans ce travail sont homogènes en termes d'IP-OMS; le caractère hautement significatif de certains aspects de la qdv dans chacune de ces études témoignant de la valeur ajoutée des scores de qdv par rapport à l'IP-OMS et aux paramètres clinico-biologiques. Cela suggère la nécessité d'une auto-évaluation de la qdv du patient au moment du diagnostic de cancer en plus de l'IP-OMS (évalué par le clinicien) d'autant plus que Bottomley [86] a rapporté une sous-estimation de l'état fonctionnel et une sous-estimation des symptômes du patient par le clinicien.

La qdv pourrait donc améliorer la communication patient/médecin comme cela a été établi dans une petite étude comparant la consultation de patients atteints de cancer [87] avec ou sans auto-évaluation de la qdv en salle d'attente. En plus de l'amélioration de la communication, l'étude a montré une plus grande satisfaction du patient du médecin.

De façon générale, les scores de qdv ont permis d'améliorer les performances des systèmes de classification dans le cancer colorectal métastatique (systèmes de Köhne et du GERCOR) mais aussi dans le CHC avancé après détermination des valeurs seuil optimaux (systèmes BCLC, CLIP, GRETCH et BoBar). Les valeurs de C-index observées pour les modèles avec et sans données de qdv le confirment. Le C-index varie de 0.65 à 0.67 pour le modèle multivarié dans le CCR mais les différences sont plus importantes pour l'amélioration des systèmes Köhne (C-index varie de 0.54 à 0.67 mais LDH dans le modèle) et GERCOR (0.63 à 0.68 grâce à 2 dimensions de la qdv uniquement). A noter que le gain en C-index pour le modèle de Köhne est plus important dans notre étude que dans l'étude d'Efficace et al.[43] (0.620 à 0.648). Cependant Efficace et al.[43] ont utilisé les variables brutes du score

de Köhne alors que le score de Köhne avec trois groupes pronostiques est utilisé notre ce travail ; cela pourrait expliquer en partie le gain plus important de C-index que nous avons observé. Pour le CHC, après l'ajout de la qdv aux scores existants, les différences de C-index vont de 0.04 (BCLC) à 0.05 (GRETCH, CLIP et BoBar). Si l'on tient compte des paramètres clinico-biologiques en plus de la qdv, les C-index sont plus élevés et le BCLC est le système qui a le plus gagné en C-index (de 0.58 à 0.68) confirmant sa faible performance pour les patients atteints de CHC avancé. Tous les NRI et IDI sont pratiquement significativement différents de zéro sauf pour le score BoBar. Ces gains sont souvent plus importants à 3 mois et décroissent avec le temps. Cela suggère que ces scores pronostiques révisés sont plus pertinents pour identifier les patients de très mauvais pronostic, qui ne devraient pas être inclus dans les essais cliniques. Ce dernier résultat s'il est confirmé, pourrait s'ajouter à la liste des causes d'échec des phases III dans le CHC décrites par Llovet et al. [79]. En effet l'inclusion dans le bras expérimental de patients avec une mauvaise qdv pourrait augmenter le risque de décès à court terme. Cependant une validation externe est indispensable avant l'implémentation de nos résultats en routine clinique. Dans le cadre de l'ACP, l'intérêt de la qdv est confirmé par la classification du patient en bon ou mauvais pronostic grâce au seul score de fatigue.

La différence de valeurs seuil pour la fatigue entre le CHC (66.67) et le cancer du pancréas (47) pourrait s'expliquer par le fait que le CHC se développe souvent sur un foie cirrhotique et que ces patients ont probablement une évolution de la perception de leur santé (« response shift »[88]). Au contraire, ce phénomène de « response shift » est moins probable dans le cancer du pancréas métastatique dont l'évolution est généralement très rapide.

L'exclusion des patients dont la douleur n'est pas contrôlée de l'étude FIRGEM (cancer du pancréas métastatique) ainsi que la différence de méthodologie utilisée pour la recherche de valeurs seuil pourraient également expliquer la différence de valeurs seuil observée.

La non-uniformité des outils pour évaluer la qdv (EQ-5D pour le cancer colorectal et QLQ-C30 pour le CHC et pancréas) est une difficulté dans ce travail car l'utilisation d'un même outil pour toutes les localisations aurait permis une interprétation plus simple des résultats ainsi qu'une comparaison entre les différents types de cancer. Cela aurait également permis une diffusion plus aisée du message auprès des cliniciens et aurait ainsi facilité une standardisation des pratiques. A l'inverse la

concordance des résultats de différents questionnaires sur la valeur pronostique de la qdv confirme son intérêt pronostique dans le cancer.

Le fait que les données utilisées dans ce travail proviennent d'essais cliniques et non pas de la routine clinique est également une limite à la généralisation des résultats pour l'ensemble des patients atteints des trois types de cancer étudiés. En effet, seuls les patients ayant certains critères d'inclusion et de non inclusion (donc sur-sélectionnés) participent aux essais cliniques.

Même si l'algorithme d'imputation des données manquantes semble performant dans le cadre des forêts aléatoires appliquées aux données sur le cancer du pancréas, la forte proportion de données manquantes de qdv ainsi que dans le cas du cancer colorectal métastatique est une autre limite de ce travail.

VI. Perspectives :

En perspective, plus de travaux collaboratifs entre les acteurs des sciences sociales et les épidémiologistes serait important pour mieux comprendre le mécanisme par lequel la qdv influence indépendamment des facteurs clinico-biologiques le pronostic des patients atteints de cancer. En effet, la qdv est influencée par les stratégies individuelles d'adaptation du patient (stratégies de coping) propre à chaque individu ; ces stratégies d'adaptation étant elles-mêmes dépendantes de la personnalité, de la culture, de la spiritualité etc., la qdv pourrait donc n'être qu'un médiateur de l'effet de ces différentes caractéristiques du patient.

Etant donnée la diversité des questionnaires utilisés dans les études concernant la valeur pronostique de la qdv, davantage d'études comparatives (utilisant la même méthodologie statistique) de ces questionnaires quant à leur sensibilité à détecter les dimensions pertinentes pour prédire le pronostic devraient être réalisées pour émettre des recommandations pour le questionnaire à utiliser. Une standardisation des outils de mesures de qdv serait alors possible facilitant ainsi la prise en compte des données de qdv dans le quotidien des cliniciens.

Un travail concernant les valeurs seuil des scores de qdv dans d'autres localisations tumorales autres que le CHC et l'ACP pourrait faciliter son utilisation par les cliniciens pour améliorer la communication patient/médecin et pour adapter la prise en charge en prenant en compte le ressenti du patient. Ces valeurs seuil seraient également importantes dans la planification des essais cliniques.

L'intérêt de la qdv pour prédire la réponse au traitement ou la rechute devrait également être plus étudié dans le but d'une éventuelle meilleure personnalisation de la prise en charge basée sur les mesures de qdv.

Malgré leur apparente facilité d'interprétation, l'utilité des indices NRI et IDI pour évaluer l'intérêt d'un nouveau marqueur n'est pas encore prouvée[89]. Davantage de recherche est nécessaire pour étudier l'apport de ces indices NRI et IDI aux mesures traditionnelles les plus utilisées que sont le c-index, la statistique de Schemper et le test de monotonicité du gradient.

La méthode de Faraggi s'est avérée très intéressante pour la recherche des valeurs seuil. De nouvelles études pour confirmer nos valeurs seuil dans le CHC sont nécessaires. La méthode de Faraggi pourrait également être utilisée pour d'autres localisations cancéreuses pour déterminer des valeurs seuil pour les scores de qdv.

VII. Conclusion :

Ce travail confirme la valeur pronostique de la qdv dans le CCRI métastatique, l'ACP métastatique et le CHC avancé. L'ajout des scores de qdv a permis d'améliorer la performance de tous les systèmes de classification étudiés.

Les valeurs seuil trouvées pour le score de fatigue dans l'adénocarcinome du pancréas et pour les scores de fatigue, diarrhée et bien-être physique dans le CHC devraient permettre une utilisation plus facile des informations provenant de la qdv du patient lors de la discussion patient/médecin ainsi que dans la prise en charge des patients atteints de ces types de cancer. Ces valeurs seuils pourraient également servir dans la planification des essais cliniques en tant que critères d'inclusion/non inclusion ou de stratification.

VIII. Bibliographie :

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, et al. (2014) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer J Int Cancer. doi:10.1002/ijc.29210.

2. Slevin ML, Plant H, Lynch D, Drinkwater J, Gregory WM (1988) Who should measure quality of life, the doctor or the patient? Br J Cancer 57: 109–112.

3. Quinten C, Coens C, Mauer M, Comte S, Sprangers MAG, et al. (2009) Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. Lancet Oncol 10: 865–871. doi:10.1016/S1470-2045(09)70200-1.

4. Reck M, Thatcher N, Smit EF, Lorigan P, Szutowicz-Zielinska E, et al. (2012) Baseline quality of life and performance status as prognostic factors in patients with extensive-stage disease small cell lung cancer treated with pemetrexed plus carboplatin vs. etoposide plus carboplatin. Lung Cancer Amst Neth 78: 276–281. doi:10.1016/j.lungcan.2012.09.002.

5. Tournigand C, Cervantes A, Figer A, Lledo G, Flesch M, et al. (2006) OPTIMOX1: a randomized study of FOLFOX4 or FOLFOX7 with oxaliplatin in a stop-and-Go fashion in advanced colorectal cancer--a GERCOR study. J Clin Oncol Off J Am Soc Clin Oncol 24: 394–400. doi:10.1200/JCO.2005.03.0106.

6. Barbare J-C, Bouché O, Bonnetain F, Dahan L, Lombard-Bohas C, et al. (2009) Treatment of advanced hepatocellular carcinoma with long-acting octreotide: a phase III multicentre, randomised, double blind placebo-controlled study. Eur J Cancer Oxf Engl 1990 45: 1788–1797. doi:10.1016/j.ejca.2009.02.018.

7. Trouilloud I, Dupont-Gossard A-C, Malka D, Artru P, Gauthier M, et al. (n.d.) Fixed-dose rate gemcitabine alone or alternating with FOLFIRI.3 (irinotecan, leucovorin and fluorouracil) in the first-line treatment of patients with metastatic pancreatic adenocarcinoma: An AGEO randomised phase II study (FIRGEM). Eur J Cancer 0. Available: http://www.ejcancer.com/article/S0959804914009988/abstract. Accessed 31 October 2014.

8. Köhne CH, Cunningham D, Di Costanzo F, Glimelius B, Blijham G, et al. (2002) Clinical determinants of survival in patients with 5-fluorouracil-based treatment for metastatic colorectal cancer: results of a multivariate analysis of 3825 patients. Ann Oncol Off J Eur Soc Med Oncol ESMO 13: 308–317.

9. Chibaudel B, Bonnetain F, Tournigand C, Bengrine-Lefevre L, Teixeira L, et al. (2011) Simplified prognostic model in patients with oxaliplatin-based or irinotecan-based first-line chemotherapy for metastatic colorectal cancer: a GERCOR study. The Oncologist 16: 1228–1238. doi:10.1634/theoncologist.2011-0039.

10. Carlens E, Dahlström G, Nöu E (1971) An attempt to include “quality of life” in evaluation of survival in bronchial cancer therapy. Bronches 21: 215–219.

11. Burge PS, Prankerd TA, Richards JD, Sare M, Thompson DS, et al. (1975) Quality and quantity of survival in acute myeloid leukaemia. Lancet 2: 621–624.

12. Gilbert MR, Dignam JJ, Armstrong TS, Wefel JS, Blumenthal DT, et al. (2014) A randomized trial of bevacizumab for newly diagnosed glioblastoma. N Engl J Med 370: 699–708. doi:10.1056/NEJMoa1308573.

13. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, et al. (2005) Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. N Engl J Med 352: 987–996. doi:10.1056/NEJMoa043330.

14. Pater JL, Loeb M (1982) Nonanatomic prognostic factors in carcinoma of the lung: a multivariate analysis. Cancer 50: 326–331.

15. Montazeri A (2009) Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. Health Qual Life Outcomes 7: 102. doi:10.1186/1477-7525-7-102.

16. Jacot W, Colinet B, Bertrand D, Lacombe S, Bozonnat M-C, et al. (2008) Quality of life and comorbidity score as prognostic determinants in non-small-cell lung cancer patients. Ann Oncol Off J Eur Soc Med Oncol ESMO 19: 1458–1464. doi:10.1093/annonc/mdn064.

17. Ware JE, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30: 473–483.

18. EuroQol Group (1990) EuroQol--a new facility for the measurement of health-related quality of life. Health Policy Amst Neth 16: 199–208.

19. Spitzer WO, Dobson AJ, Hall J, Chesterman E, Levi J, et al. (1981) Measuring the quality of life of cancer patients: a concise QL-index for use by physicians. J Chronic Dis 34: 585–597.

20. Schipper H, Clinch J, McMurray A, Levitt M (1984) Measuring the quality of life of cancer patients: the Functional Living Index-Cancer: development and validation. J Clin Oncol 2: 472–483.

21. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, et al. (1993) The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol Off J Am Soc Clin Oncol 11: 570–579.

22. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, et al. (1993) 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 85: 365–376.

23. Ward WL, Hahn EA, Mo F, Hernandez L, Tulsky DS, et al. (1999) Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. Qual Life Res 8: 181–195. doi:10.1023/A:1008821826499.

24. Whistance RN, Conroy T, Chie W, Costantini A, Sezer O, et al. (2009) Clinical and psychometric validation of the EORTC QLQ-CR29 questionnaire module to assess health-related quality of life in patients with colorectal cancer. Eur J Cancer Oxf Engl 1990 45: 3017–3026. doi:10.1016/j.ejca.2009.08.014.

25. Blazeby JM, Currie E, Zee BCY, Chie W-C, Poon RT, et al. (2004) Development of a questionnaire module to supplement the EORTC QLQ-C30 to assess quality of life in patients with hepatocellular carcinoma, the EORTC QLQ-

HCC18. Eur J Cancer Oxf Engl 1990 40: 2439–2444. doi:10.1016/j.ejca.2004.06.033.

26. Fitzsimmons D, Johnson CD, George S, Payne S, Sandberg AA, et al. (1999) Development of a disease specific quality of life (QoL) questionnaire module to supplement the EORTC core cancer QoL questionnaire, the QLQ-C30 in patients with pancreatic cancer. EORTC Study Group on Quality of Life. Eur J Cancer Oxf Engl 1990 35: 939–941.

27. Heffernan N, Cella D, Webster K, Odom L, Martone M, et al. (2002) Measuring Health-Related Quality of Life in Patients With Hepatobiliary Cancers: The Functional Assessment of Cancer Therapy–Hepatobiliary Questionnaire. J Clin Oncol 20: 2229–2239. doi:10.1200/JCO.2002.07.093.

28. Wisløff F, Hjorth M (1997) Health-related quality of life assessed before and during chemotherapy predicts for survival in multiple myeloma. Nordic Myeloma Study Group. Br J Haematol 97: 29–37.

29. Efficace F, Biganzoli L, Piccart M, Coens C, Van Steen K, et al. (2004) Baseline health-related quality-of-life data as prognostic factors in a phase III multicentre study of women with metastatic breast cancer. Eur J Cancer Oxf Engl 1990 40: 1021–1030. doi:10.1016/j.ejca.2004.01.014.

30. Montazeri A, Milroy R, Hole D, McEwen J, Gillis CR (2001) Quality of life in lung cancer patients: as an important prognostic factor. Lung Cancer Amst Neth 31: 233–240.

31. Roychowdhury DF, Hayden A, Liepa AM (2003) Health-related quality-of-life parameters as independent prognostic factors in advanced or metastatic bladder cancer. J Clin Oncol Off J Am Soc Clin Oncol 21: 673–678.

32. Coyne JC, Pajak TF, Harris J, Konski A, Movsas B, et al. (2007) Emotional well-being does not predict survival in head and neck cancer patients: a Radiation Therapy Oncology Group study. Cancer 110: 2568–2575. doi:10.1002/cncr.23080.

33. Luoma M-L, Hakamies-Blomqvist L, Sjöström J, Pluzanska A, Ottoson S, et al. (2003) Prognostic value of quality of life scores for time to progression (TTP) and overall survival time (OS) in advanced breast cancer. Eur J Cancer Oxf Engl 1990 39: 1370–1376.

34. Cella D, Bushmakin AG, Cappelleri JC, Charbonneau C, Michaelson MD, et al. (2012) Baseline quality of life as a prognostic survival tool in patients receiving sunitinib for metastatic renal cell carcinoma. Br J Cancer 106: 646–650. doi:10.1038/bjc.2011.589.

35. Quinten C, Martinelli F, Coens C, Sprangers MAG, Ringash J, et al. (2014) A global analysis of multitrial data investigating quality of life and symptoms as prognostic factors for survival in different tumor sites. Cancer 120: 302–311. doi:10.1002/cncr.28382.

36. Field K, Lipton L (2007) Metastatic colorectal cancer-past, progress and future. World J Gastroenterol WJG 13: 3806–3815.

37. Folprecht G, Grothey A, Alberts S, Raab H-R, Köhne C-H (2005) Neoadjuvant treatment of unresectable colorectal liver metastases: correlation between tumour response and resection rates. Ann Oncol Off J Eur Soc Med Oncol ESMO 16: 1311–1319. doi:10.1093/annonc/mdi246.

38. Mitry E, Rollot F, Jooste V, Guiu B, Lepage C, et al. (2013) Improvement in survival of metastatic colorectal cancer: are the benefits of clinical trials reproduced in population-based studies? Eur J Cancer Oxf Engl 1990 49: 2919–2925. doi:10.1016/j.ejca.2013.04.001.

39. Ye L-C, Liu T-S, Ren L, Wei Y, Zhu D-X, et al. (2013) Randomized controlled trial of cetuximab plus chemotherapy for patients with KRAS wild-type unresectable colorectal liver-limited metastases. J Clin Oncol Off J Am Soc Clin Oncol 31: 1931–1938. doi:10.1200/JCO.2012.44.8308.

40. Braun DP, Gupta D, Grutsch JF, Staren ED (2011) Can changes in health related quality of life scores predict survival in stages III and IV colorectal cancer? Health Qual Life Outcomes 9: 62. doi:10.1186/1477-7525-9-62.

41. Efficace F, Bottomley A, Coens C, Van Steen K, Conroy T, et al. (2006) Does a patient's self-reported health-related quality of life predict survival beyond key biomedical data in advanced colorectal cancer? Eur J Cancer Oxf Engl 1990 42: 42–49. doi:10.1016/j.ejca.2005.07.025.

42. Maisey NR, Norman A, Watson M, Allen MJ, Hill ME, et al. (2002) Baseline quality of life predicts survival in patients with advanced colorectal cancer. Eur J Cancer Oxf Engl 1990 38: 1351–1357.

43. Efficace F, Innominato PF, Bjarnason G, Coens C, Humblet Y, et al. (2008) Validation of Patient's Self-Reported Social Functioning As an Independent Prognostic Factor for Survival in Metastatic Colorectal Cancer Patients: Results of an International Study by the Chronotherapy Group of the European Organisation for Research and Treatment of Cancer. J Clin Oncol 26: 2020–2026. doi:10.1200/JCO.2007.12.3117.

44. Llovet JM, Burroughs A, Bruix J (2003) Hepatocellular carcinoma. Lancet 362: 1907–1917. doi:10.1016/S0140-6736(03)14964-1.

45. Yeo W (2006) Quality of life is predictive of survival in patients with unresectable hepatocellular carcinoma. Ann Oncol 17: 1083–1089. doi:10.1093/annonc/mdl065.

46. Bonnetain F, Paoletti X, Collette S, Doffoel M, Bouché O, et al. (2008) Quality of life as a prognostic factor of overall survival in patients with advanced hepatocellular carcinoma: results from two French clinical trials. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 17: 831–843. doi:10.1007/s11136-008-9365-y.

47. Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, et al. (2011) FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med 364: 1817–1825. doi:10.1056/NEJMoa1011923.

48. Lis CG, Gupta D, Grutsch JF (2006) Patient satisfaction with quality of life as a predictor of survival in pancreatic cancer. Int J Gastrointest Cancer 37: 35–44. doi:10.1385/IJGC:37:1:35.

49. Ferrans CE, Powers MJ (1985) Quality of life index: development and psychometric properties. ANS Adv Nurs Sci 8: 15–24.

50. Robinson DW, Eisenberg DF, Cella D, Zhao N, de Boer C, et al. (2008) The prognostic significance of patient-reported outcomes in pancreatic cancer cachexia. J Support Oncol 6: 283–290.

51. Bernhard J, Dietrich D, Glimelius B, Hess V, Bodoky G, et al. (2010) Estimating prognosis and palliation based on tumour marker CA 19-9 and quality of life indicators in patients with advanced pancreatic cancer receiving chemotherapy. Br J Cancer 103: 1318–1324. doi:10.1038/sj.bjc.6605929.

52. Gourgou-Bourgade S, Bascoul-Mollevi C, Desseigne F, Ychou M, Bouché O, et al. (2013) Impact of FOLFIRINOX compared with gemcitabine on quality of life in patients with metastatic pancreatic cancer: results from the PRODIGE 4/ACCORD 11 randomized trial. J Clin Oncol Off J Am Soc Clin Oncol 31: 23–29. doi:10.1200/JCO.2012.44.4869.

53. Braun DP, Gupta D, Staren ED (2013) Longitudinal health-related quality of life assessment implications for prognosis in stage IV pancreatic cancer. Pancreas 42: 254–259. doi:10.1097/MPA.0b013e31825b9f56.

54. Cox DR (2005) Selected Statistical Papers of Sir David Cox: Volume 1, Design of Investigations, Statistical Methods and Applications. Cambridge University Press. 620 p.

55. Harrell FE (2001) Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer Science & Business Media. 608 p.

56. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS (2008) Random survival forests. Ann Appl Stat 2: 841–860. doi:10.1214/08-AOAS169.

57. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49: 1373–1379.

58. Vittinghoff E, McCulloch CE (2007) Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 165: 710–718. doi:10.1093/aje/kwk052.

59. Schemper M (2003) Predictive accuracy and explained variation. Stat Med 22: 2299–2308. doi:10.1002/sim.1486.

60. Pencina MJ, D'Agostino RB, Steyerberg EW (2011) Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 30: 11–21. doi:10.1002/sim.4085.

61. Van Buuren S, Boshuizen HC, Knook DL (1999) Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 18: 681–694.

62. Clark TG, Altman DG (2003) Developing a prognostic model in the presence of missing data: an ovarian cancer case study. J Clin Epidemiol 56: 28–37.

63. Mauer M, Bottomley A, Coens C, Gotay C (2008) Prognostic factor analysis of health-related quality of life data in cancer: a statistical methodological evaluation. Expert Rev Pharmacoecon Outcomes Res 8: 179–196. doi:10.1586/14737167.8.2.179.

64. Altman DG, Vergouwe Y, Royston P, Moons KGM (2009) Prognosis and prognostic research: validating a prognostic model. BMJ 338: b605.

65. Picard RR, Cook RD (1984) Cross-Validation of Regression Models. J Am Stat Assoc 79: 575–583. doi:10.1080/01621459.1984.10478083.

66. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, et al. (2014) Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiol Camb Mass 25: 114–121. doi:10.1097/EDE.0000000000000018.

67. Kishiki T, Masaki T, Matsuoka H, Kobayashi T, Suzuki Y, et al. (2013) Modified Glasgow prognostic score in patients with incurable stage IV colorectal cancer. Am J Surg 206: 234–240. doi:10.1016/j.amjsurg.2012.07.051.

68. Kobayashi H, Kotake K, Sugihara K (2013) Prognostic scoring system for stage IV colorectal cancer: is the AJCC sub-classification of stage IV colorectal cancer appropriate? Int J Clin Oncol 18: 696–703. doi:10.1007/s10147-012-0433-5.

69. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, et al. (1985) Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer 56: 918–928.

70. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators (1998). Hepatol Baltim Md 28: 751–755. doi:10.1002/hep.510280322.

71. Llovet JM, Brú C, Bruix J (1999) Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 19: 329–338. doi:10.1055/s-2007-1007122.

72. Chevret S, Trinchet JC, Mathieu D, Rached AA, Beaugrand M, et al. (1999) A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d'Etude et de Traitement du Carcinome Hépatocellulaire. J Hepatol 31: 133–141.

73. Kinoshita A, Onoda H, Imai N, Iwaku A, Oishi M, et al. (2013) The Glasgow Prognostic Score, an inflammation based prognostic score, predicts survival in patients with hepatocellular carcinoma. BMC Cancer 13: 52. doi:10.1186/1471-2407-13-52.

74. Leung TWT, Tang AMY, Zee B, Lau WY, Lai PBS, et al. (2002) Construction of the Chinese University Prognostic Index for hepatocellular carcinoma and comparison with the TNM staging system, the Okuda staging system, and the Cancer of the Liver Italian Program staging system: a study based on 926 patients. Cancer 94: 1760–1769.

75. Kudo M, Chung H, Osaki Y (2003) Prognostic staging system for hepatocellular carcinoma (CLIP score): its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging Score (JIS score). J Gastroenterol 38: 207–215. doi:10.1007/s005350300038.

76. Yau T, Tang VYF, Yao T-J, Fan S-T, Lo C-M, et al. (2014) Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. Gastroenterology 146: 1691–1700.e3. doi:10.1053/j.gastro.2014.02.032.

77. Hamada T, Nakai Y, Yasunaga H, Isayama H, Matsui H, et al. (2014) Prognostic nomogram for nonresectable pancreatic cancer treated with gemcitabine-based chemotherapy. Br J Cancer 110: 1943–1949. doi:10.1038/bjc.2014.131.

78. Hsu C-Y, Lee Y-H, Hsia C-Y, Huang Y-H, Su C-W, et al. (2013) Performance status in patients with hepatocellular carcinoma: determinants, prognostic impact, and ability to improve the Barcelona Clinic Liver Cancer system. Hepatol Baltim Md 57: 112–119. doi:10.1002/hep.25950.

79. Llovet JM, Hernandez-Gea V (2014) Hepatocellular carcinoma: reasons for phase III failure and novel perspectives on trial design. Clin Cancer Res Off J Am Assoc Cancer Res 20: 2072–2079. doi:10.1158/1078-0432.CCR-13-0547.

80. Siddiqui F, Pajak TF, Watkins-Bruner D, Konski AA, Coyne JC, et al. (2008) Pretreatment quality of life predicts for locoregional control in head and neck cancer patients: a radiation therapy oncology group analysis. Int J Radiat Oncol Biol Phys 70: 353–360. doi:10.1016/j.ijrobp.2007.06.024.

81. Kenne Sarenmalm E, Odén A, Ohlén J, Gaston-Johansson F, Holmberg SB (2009) Changes in health-related quality of life may predict recurrent breast cancer. Eur J Oncol Nurs Off J Eur Oncol Nurs Soc 13: 323–329. doi:10.1016/j.ejon.2009.05.002.

82. Braun DP, Gupta D, Staren ED (2012) Predicting survival in prostate cancer: the role of quality of life assessment. Support Care Cancer Off J Multinatl Assoc Support Care Cancer 20: 1267–1274. doi:10.1007/s00520-011-1213-x.

83. Gupta D, Braun DP, Staren ED (2013) Prognostic value of changes in quality of life scores in prostate cancer. BMC Urol 13: 32. doi:10.1186/1471-2490-13-32.

84. Svensson H, Hatschek T, Johansson H, Einbeigi Z, Brandberg Y (2012) Health-related quality of life as prognostic factor for response, progression-free survival,

and survival in women with metastatic breast cancer. Med Oncol Northwood Lond Engl 29: 432–438. doi:10.1007/s12032-011-9844-9.

85. Tsai W-L, Chien C-Y, Huang H-Y, Liao K-C, Fang F-M (2013) Prognostic value of quality of life measured after treatment on subsequent survival in patients with nasopharyngeal carcinoma. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 22: 715–723. doi:10.1007/s11136-012-0213-8.

86. Bottomley A (2002) The Cancer Patient and Quality of Life. The Oncologist 7: 120–125. doi:10.1634/theoncologist.7-2-120.

87. Detmar SB, Aaronson NK (1998) Quality of life assessment in daily clinical oncology practice: a feasibility study. Eur J Cancer Oxf Engl 1990 34: 1181–1186.

88. Hamidou Z, Dabakuyo TS, Bonnetain F (2011) Impact of response shift on longitudinal quality-of-life assessment in cancer clinical trials. Expert Rev Pharmacoecon Outcomes Res 11: 549–559. doi:10.1586/erp.11.57.

89. Hilden J, Gerds TA (2014) A note on the evaluation of novel biomarkers: do not rely on integrated discrimination improvement and net reclassification index. Stat Med 33: 3405–3414. doi:10.1002/sim.5804.