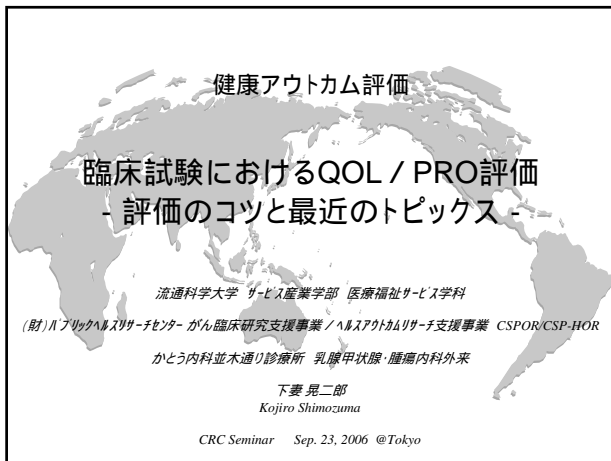



健康アウトカム評価



**臨床試験におけるQOL / PRO評価  
- 評価のコツと最近のトピックス -**

流通科学大学 サービス産業学部 医療福祉サービス学科  
(財)ハブリンクヘルスリサーチセンター がん臨床研究支援事業 / ヘルスアウトカムリサーチ支援事業 CSPOR/CSP-HOR  
かとう内科並木通り診療所 乳腺甲状腺・腫瘍内科外来  
下妻 晃二郎  
Kojiro Shimozuma  
CRC Seminar Sep. 23, 2006 @Tokyo

## 本日の話の内容



- 本日は応用編が中心 (一部前回と重複)
  - 前は基礎編が中心
- 基礎
  - 「患者報告に基づくアウトカム: Patient-reported outcomes (PRO)評価の意義
- 応用
  - QOL/PROガイダンスの紹介
  - QOL/PRO調査のコツ
  - QOL/PRO解析における課題

### プロフィール型QOL尺度(例) QOL-ACD

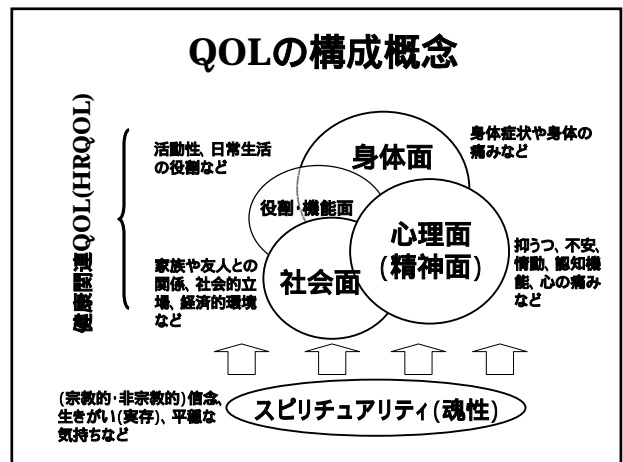
がん薬物療法における QOL 調査票 (QOL-ACD)

氏名: \_\_\_\_\_ 平成 \_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日  
年齢: \_\_\_\_\_ 歳 性別: 1. 男 2. 女 体重: \_\_\_\_\_ Kg

この調査票は、あなたの現在の状態を正しく理解するために用いるものです。  
ここ数日間のあなたの状態にあてはまると思われる番号に○をつけてください。  
(個人のプライバシーが外部にもれたり、治療のうえで不利益になることは決してありませんので、感じたありのままをお答え下さい。)

(この数日の間)

1. 日常生活 (活動) ができましたか。……………	1	2	3	4	5
	全くできなかった				十分できた
2. ひとりで外出することができましたか。……………	1	2	3	4	5
	全くできなかった				十分できた
3. 30分くらいの散歩はできましたか。……………	1	2	3	4	5
	全くできなかった				十分できた
4. 少し歩いてもらいたいと思いましたが。……………	5	4	3	2	1
	全く問題なかった				非常につらかった
5. 階段の昇り降りができましたか。……………	1	2	3	4	5
	全くできなかった				十分できた



### QOL-ACDの構成

尺度	略語	項目番号
1 活動性	FWB	1 to 6
2 身体状況	PWB	7 to 11
3 精神・心理状態	EWB	12 to 16
4 社会性	SWB	17 to 21
5 全般的QOL	GQL	22

- 肺がん臨床試験におけるQOL評価
  - 解析結果の例 (Fig.)
    - Copyrightの制約のため表示せず

- DC is significantly better than VdsC in Physical aspect of QOL.

価値付け型尺度(例) EQ-5D

Copyrightの制約のため表示せず

表3 日本語版EQ-5Dの効用値換算表

Copyrightの制約のため表示せず

池上直己 他:臨床のためのQOL評価ハンドブック 医学書院 2001 p16

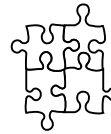
本日の話の内容



- 基礎
  - 「患者報告に基づくアウトカム: Patient-reported outcomes (PRO)評価の意義
- 応用
  - QOL/PROガイダンスの紹介
  - QOL/PRO調査のコツ
  - QOL/PRO解析における課題

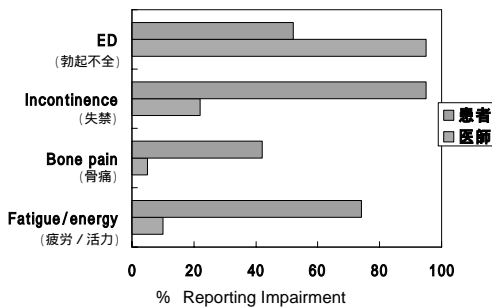
QOLの2つの特徴

多要素性  
Multidimensionality  
(Multidomain concept)



主観性  
Subjectivity

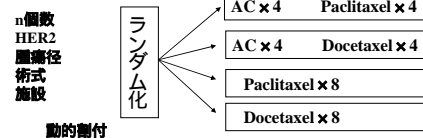
医師の評価と患者の評価の乖離例(1)

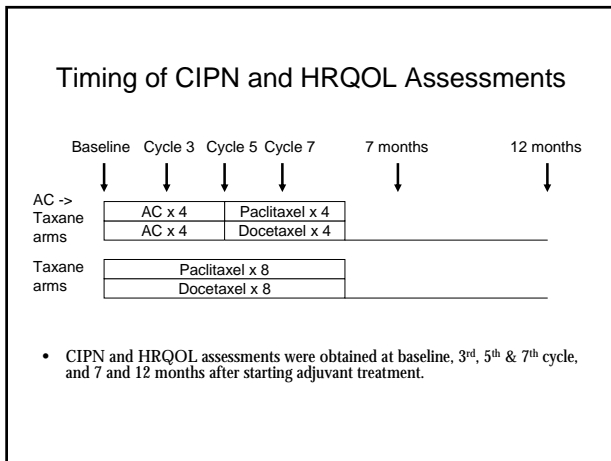


Litwin MS, et al: J Urol, 59:1988-92, 1998  
京大 福原俊一先生スライドより一部改変

医師の評価と患者の評価の乖離例(2)

- 末梢神経障害(抗がん剤の副作用)に関する乖離
- N-SAS BC 02 (2001-2003)
- 腋窩リンパ節転移陽性乳癌患者を対象とした術後補助療法臨床試験 (N=1200)
- エンドポイント(評価指標)
  - 無病生存期間
  - 神経毒性を中心とした健康関連QOL





### Chemotherapy-Induced Peripheral Neuropathy (CIPN) and HRQOL Instruments

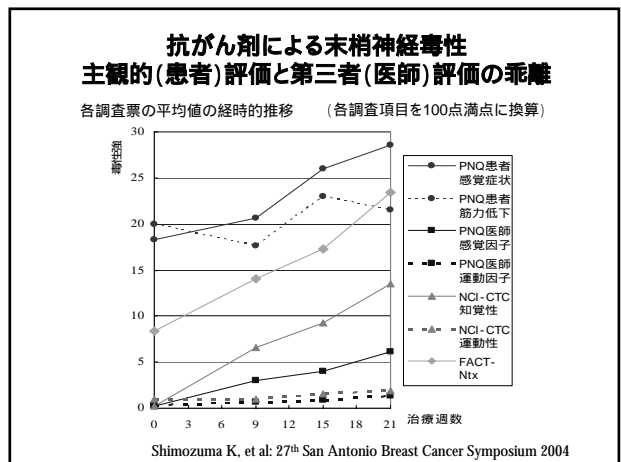
Patient-based	Physician-based
<b>PNQ*</b> (Patient Neurotoxicity Questionnaire)	<b>NCI-CTC (Version 2.0)</b> (Neuro-sensory & Neuro-motor)
<b>FACT/GOG-Ntx**</b> (Functional Assessment of Cancer Therapy-Neurotoxicity)	<b>PNEF***</b> (Physician Neurotoxicity Examination Form)

\* Hausheer, et al., Semin Oncol 2006, Shimozuma, et al., 2004 SABCS  
 \*\* Calhoun, et al., Int J Gynecol Cancer 2003  
 \*\*\* Developed by Hausheer, et al

### Patient Neurotoxicity Questionnaire (PNQ)

Item	A	B	C	D*	E*
Item 1	I have no numbness, pain, or tingling in my hands or feet.	I have mild tingling, pain or numbness in my hands or feet. This does not interfere with my activities.	I have moderate tingling, pain or numbness in my hands or feet. This does not interfere with my activities of daily living.	I have moderate to severe tingling, pain or numbness in my hands or feet. This interferes with my activities of daily living.	I have severe tingling, pain or numbness in my hands or feet. It completely prevents me from doing most activities.
Item 2	I have no weakness in my arms or legs.	I have mild weakness in my arms or legs. This does not interfere with my activities.	I have moderate weakness in my arms or legs. This does not interfere with my activities of daily living.	I have moderate to severe weakness in my arms or legs. This interferes with my activities of daily living.	I have severe weakness in my arms or legs. It completely prevents me from doing most activities.

\*Patients answering D or E provided additional information on specific Activities of Daily Living (ADL's) that were affected.



### Sensitivity

- Distribution of the scores (PNQ vs. NCI-CTC) -

PNQ-sensory	NCI-CTC-sensory					PNQ-motor	NCI-CTC-motor				
	0	1	2	3	4		0	1	2	3	4
A(0)	489	38	0	0	0	A(0)	492	5	1	0	0
B(1)	432	252	4	0	0	B(1)	701	37	2	0	0
C(2)	113	171	5	0	0	C(2)	231	17	3	0	0
D(3)	44	66	11	3	1	D(3)	62	10	5	1	0
E(4)	9	1	0	0	0	E(4)	9	0	0	0	0

Kappa=0.16 (sensory), Kappa=0.02 (motor)

- PNQ scores distributed between A and E (full range), while NCI-CTC scores mainly distributed between 0 and 1.
- Sensitivity of NCI-CTC, when PNQ scores are regarded as to be true.
  - AB/CDE and 01/234: Sensitivity=0.047 (sensory), 0.027 (motor)
  - ABC/DE and 012/34: Sensitivity=0.030 (sensory), 0.011 (motor)
- Specificity distributed between 0.996 and 1.

Shimozuma K, et al: 2006 ASCO Annual Meeting

- ### Assessing health-related quality of life in palliative care: Comparing patient and physician assessments
- (Petersen MA, et al. Eur J Cancer 2006)
- 目的
    - 医師評価と患者評価の比較
  - 対象
    - がん緩和医療中353人(デンマーク)
  - 使用尺度
    - がん特異的QOL尺度 EORTC QLQ-C30
  - 調査ポイント
    - 13週にわたり毎週(可能な限り)

• EORTC QLQ-C30 ver. 3  
 - Copyrightの制約のため表示せず

### EORTC QLQ-C30の構成

Scale	Item
<b>Global health status / QoL 全般的健康</b>	
Global health status/QoL 全般的健康	QL 29, 30
<b>Functioning scales 機能尺度</b>	
Physical functioning 身体機能	PF 1 to 5
Role functioning 役割機能	RF 6, 7
Emotional functioning 心理機能	EF 21 to 24
Cognitive functioning 認知機能	CF 20, 25
Social functioning 社会機能	SF 26, 27
<b>Symptom scales and/or items 症状尺度 / 項目</b>	
Fatigue 倦怠感	FA 10, 12, 18
Nausea and vomiting 悪心・嘔吐	NV 14, 15
Pain 疼痛	PA 9, 19
Dyspnoea 呼吸困難感	DY 8
Insomnia 不眠	SL 11
Appetite loss 食欲不振	AP 13
Constipation 便秘	CO 16
Diarrhoea 下痢	DI 17
Financial difficulties 経済的困難	FI 28

### 医師 - 患者間評価の一致性

Variable	Correlation (Spearman)	Agreement (weighted Kappa)
<b>Functioning scale</b>		
Physical (PF)	0.74	0.47
Emotional (EF)	0.28	0.15
Cognitive (CF)	0.46	0.24
Social (SF)	0.23	0.15
<b>Symptom scale</b>		
Fatigue (FA)	0.54	0.37
Nausea/Vomiting (NV)	0.76	0.54
Pain (PA)	0.67	0.47
Dyspnea	0.76	0.51
Insomnia	0.38	0.23
Appetite loss	0.57	0.40
Constipation	0.74	0.60
Diarrhea	0.40	0.33

Highest value     lowest value

Petersen MA, et al. Eur J Cancer 2006

### 医師 - 患者間評価の一致性 - 概要 -

- 過大評価か過小評価か？
  - 一般に医師が過小評価していることが多い
  - 機能スケール 2/4, 症状スケール 8/8 で過小評価
- 一致性
  - 相関係数 (Spearman)
    - 0.23 (社会機能) - 0.76(悪心・嘔吐)
  - 一致率 (weighted Kappa)
    - 0.15 (心理機能, 社会機能) - 0.60(便秘)
  - 一致の程度
    - Slight 心理機能, 社会機能
    - Fair 不眠, 食欲不振, 下痢, 倦怠感, 認知機能
    - Moderate 身体機能, 疼痛, 呼吸困難, 便秘, 悪心・嘔吐

Petersen MA, et al. Eur J Cancer 2006

Fig. 1 Patient-physician agreement  
 Copyrightの制約のため表示せず

●認知機能(CF)  
 ●心理機能(EF)

●患者の評価に関わらず水平  
 ●患者の訴えがよくなるからない？

●患者の訴えが弱いと、過大評価する傾向

Petersen MA, et al. Eur J Cancer 2006

Fig. 1 Patient-physician agreement  
 Copyrightの制約のため表示せず

●心理機能(EF)  
 ●社会機能(SF)

●患者の訴えが強くなると、過小評価が進む

Petersen MA, et al. Eur J Cancer 2006

Fig. 1 Patient-physician agreement  
Copyrightの制約のため表示せず

●症状スケール

- 全体的に過小評価傾向
- 患者の訴えが強くなると過小評価の傾向が進む

Petersen MA, et al. Eur J Cancer 2006

## 本日の話の内容

- 基礎
  - 「患者報告に基づくアウトカム: Patient-reported outcomes (PRO)評価の意義
- 応用
  - QOL/PROガイダンスの紹介
  - QOL/PRO調査のコツ
  - QOL/PRO解析における課題

### 臨床試験においてQOL / PROをエンドポイントとして用いる場合の適切なガイダンスは？

- 'Guidelines for reporting results of quality of life assessments in clinical trials'
  - Quality of Life Assessment in Clinical Trials (Staquet MJ, et al, 1998)
  - 11 pageのチェックリストのようなもの
- 'Minimum standard checklist for evaluating HRQOL outcomes in cancer clinical trials'
  - Beyond the Development of Health-Related Quality-of-Life (HRQOL) Measures: A Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials—Does HRQOL Evaluation in Prostate Cancer Research Inform Clinical Decision Making? (Efficace F, et al, J Clin Oncol 2003)
  - 次のスライド

### HRQOL評価論文(がん領域)のチェックリスト

Minimum standard checklist for evaluating HRQOL outcomes in cancer clinical trials

概念	事前に仮説(エンドポイントと期待される差)が設定されているか 尺度選定の根拠が述べられているか	Yes No N/A*
測定	尺度の計量心理的特性が報告されているか、すでに文献があるか+ 対象集団への適応の妥当性が検証されているか がん患者一般にとって、かつ/あるいは 研究目的にとって重要な領域がカバーされているか	Yes No Yes No N/A** Yes No
方法論	測定法(誰がどのような設定で)が報告されているか 事前測定のコンプライアンスが報告されているか+ 測定時期が記載されているか 欠損状況が記載されているか+	Yes No Yes No Yes No Yes No
解釈	患者の立場からの臨床的有意性の議論がなされているか 結果のいかに関わらず、結果セクション かつ/あるいは 議論のセクションで議論されているか	Yes No Yes No

\* 探索的評価と明確にうたっている場合 + 特に重要な項目  
\*\* 同様の対象集団ですでに検証されている場合  
Efficace F, et al J Clin Oncol 21:3502-3511, 2003

## 「乳癌患者のQOL評価研究のためのガイドライン」

日本乳癌学会編 2002  
[http://www.jbcs.gr.jp/member\\_o/member/QOL\\_Ver1/QOL.html](http://www.jbcs.gr.jp/member_o/member/QOL_Ver1/QOL.html)  
 member's area (ID: JBBS, Password: 2006)

## 臨床試験(治験)においてQOL / PROを評価する際の公的なガイダンスが登場

↓

欧州および米国の規制当局から発信されたガイダンス

### European Medical Agency (EMA)

医薬品評価におけるHRQOLの扱いに関するガイドンス

(2005年7月)

### European Medical Agency (EMA) ガイドンス - PROの定義 -

- Any outcome evaluated directly by the patient himself and based on patient's perception of a disease and its treatment(s) is called patient-reported outcome (PRO).
- 患者自身が直接評価した、そして病気や治療に関する患者の認識に基づくあらゆるアウトカムを、PROと呼ぶ。

### European Medical Agency (EMA) ガイドンス - 内容の概略 -

- 特徴
  - わずか5ページ。内容的には充分とはいえない。
- 目次
  - 1. Introduction (background)
    - PRO、HRQLの定義など
  - 2. HRQL in drug evaluation process
    - 尺度の選択および使い方の概略
  - 3. Study design for HRQL assessment
  - 4. Statistical analysis and hypothesis
    - エンドポイントの多重性、欠測値を避ける方法、意味のある差(MID)など

### U.S. Food and Drug Administration (FDA)

医薬品申請時のPROの取り扱いに関するガイドンス

(2006年2月  
(ドラフト))

おそらく今年末から  
来年にかけて  
正式発表

### FDAガイドンス - PROの定義 -

- A PRO is a measurement of any aspect of a patient's health status that comes directly from the patient (i.e., without the interpretation of the patients' responses by a physician or anyone else).
- PROは、患者から直接得られた、あらゆる健康面についての測定値である。
  - すなわち、医師やその他の誰の解釈も受けていない、患者の反応である。

### FDAガイドンス - PROの定義 - (つづき)

- In clinical trials, a PRO instrument can be used to measure the impact of an intervention on one or more aspects of patients' health status, hereafter referred to as PRO concepts, ranging from the purely symptomatic (response of a headache) to more complex concepts (e.g., ability to carry out activities of daily living), to extremely complex concepts such as *quality of life*, which is widely understood to be a multidomain concept with physical, psychological, and social components.
- 臨床試験では、PRO測定尺度は、ある介入が患者の健康状態の様々な側面に及ぼす影響についての測定に用いることができる。
- PROの概念は、頭痛などの純粋に症状に関するものから、日常生活を送ることができる能力、などのより複雑な概念、さらには、身体面、心理面、社会面などの多要素の概念として理解されている「QOL」のようなもっと複雑な概念までを含む。

### FDAガイダンス - 内容の概略 -

- 特徴
  - 32ページ。教科書並みに詳しい。
- 目次
  - 1. Introduction 2. Background
  - 3. PRO - regulatory perspective
  - 4. Evaluating PRO instruments
    - 概念構造、尺度開発、尺度理論など
  - 5. Study design
    - マスク、ランダム化、欠測値を避ける方法、エンドポイントの多重性など
  - 6. Data analysis
    - 多重性、欠測値、結果の解釈など
  - 用語集

### 本日の話の内容



- 基礎
  - 「患者報告に基づくアウトカム: Patient-reported outcomes (PRO)評価の意義
- 応用
  - QOL/PROガイダンスの紹介
  - QOL/PRO調査のコツ
  - QOL/PRO解析における課題

### QOL / PRO調査のコツ

- 患者さんに対する倫理性の確保
- データの科学性を保つ



このアンケートは一つ一つ大切なものです。  
あなたが自分自身で回答してください。



- 患者の「主観」を理解したい
  - 科学性(調査の目的)の確保

- 欠測データをなるべく減らしたい
  - 科学性(結果の信頼性)の確保

- 強制にならないように配慮
  - 倫理性の確保
- でも、なるべく欠測データを減らしたい
  - 科学性の確保

なるべくすべての質問に回答してください。

ただし、どうしても答えたくない質問があった場合は、その質問を飛ばして、次の質問からまた答え始めてください。



1. プライバシーは保護されています。  
2. アンケートに書かれている時間や期間に注意して回答してください。  
3. 一つの質問にあまり時間をかけず、感じたありのままを答えてください。  
4. 試験ではないので、正しい答えやまちがった答えはありません。

倫理性の確保

科学性の確保

科学性の確保

不安感の除去  
- 倫理性の確保

これらの事に注意してください。

アンケート 実地中

患者さんから、アンケートに対して質問があった場合、懇話さくめえずにアンケートの回答をいり返してください。またでも理解できない場合は、簡単に言葉を変えていり返してください。

患者さん「もういっぺん説明した時にできるだけ細かく教えてください。」

倫理性の確保

バイアスの排除  
- 科学性の確保

患者さん「アンケートは強制に行ってください。」

医師の話や表情によるバイアスの排除  
- 科学性の確保

患者さん「患者さんの健康状態や生命にたいしては心配しないでください。」

バイアスの排除  
- 科学性の確保

患者さん「付添人はアンケート中はできるだけ距離しないでください。」

付添人の話や表情によるバイアスの排除  
- 科学性の確保

アンケート回収時後すぐに、記載漏れがないかどうか確認してください。記載漏れがあった場合は、答えたくなくて答えなかった場合を除いて、再度記載を依頼してください。

バイアスの排除  
- 科学性の確保

アンケートは感謝の気持ちを持っておこなってください。

おまがいたします。

よろしく



## 本日の話の内容



- 基礎
  - 「患者報告に基づくアウトカム: Patient-reported outcomes (PRO)評価の意義
- 応用
  - QOL/PROガイダンスの紹介
  - QOL/PRO調査のコツ
  - QOL/PRO解析における課題

## GCa vs. MIC in stage IIIB or IV NSCLC

ISRCTN 52253218 London Lung Cancer Group

## Overall survival

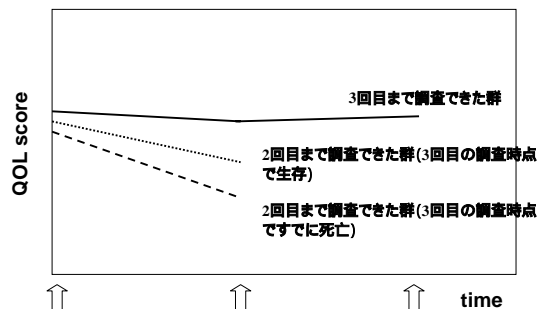
- GCa is significantly better than MIC in OS.

## Progression-free survival

- No significant difference are observed in PFS.

## 欠測値のQOLスコアへの影響

- Missing not at random (MNAR)が多い場合 -



## 解析の課題

- 欠測値(missing data)を最小限にとどめる
  - 調査票の回収率を高めるための努力
    - 臨床研究コーディネーター (CRC)の協力
    - 調査の督促体制
    - インセンティブ(研究者、患者)
- 欠測値が出る場合は、その理由調査を必ず同時に行う
  - Missing at random (MAR)か、missing not at random (MNAR)かによって解析方法が変わる
    - 特にランダム化比較試験(RCT)の場合

### 欠測値が多くある場合の解析法

- 様々なimputation (転嫁) による感度分析を行うことが多い
- 一般的な方法
  - Last value carried forward: LVCF
    - その人の最後のスコアを用いる
  - Hot deck imputation
    - その研究の他の患者のスコアをランダムに用いる
  - Cold deck imputation
    - 他の研究の結果から代入する
  - Mean imputation
  - その他、統計学的手法
    - Regression imputation
    - Multiple imputation
    - Likelihood-based approach

### 推奨参考文献

- 臨床のためのQOL評価ハンドブック
  - 池上直己, 福原俊一, 下妻晃二郎, 池田俊也, 編
  - 医学書院 ¥2,800
- QOL評価学 - 測定, 解析, 解釈のすべて
  - ビーター・M・フェイヤーズ, デビッド・マッキン著
  - 福原俊一, 数間恵子, 編
  - 中山書店 東京 2005 ¥7,980
- Quality of Life Assessment in Clinical Trials - Methods and Practice
  - Staquet MJ, Hays RD, and Fayers PM
  - Oxford University Press 約¥9,700
- 乳癌患者のQOL評価研究のためのガイドライン
  - 日本乳癌学会編 2002
  - [http://www.jbcs.gr.jp/QOL\\_Ver1/QOL.html](http://www.jbcs.gr.jp/QOL_Ver1/QOL.html)

ご静聴有難うございました!



Kojiro\_Shimozuma@red.umds.ac.jp