

CSPOR-BC新規試験
RESQ trialの紹介

岡山大学病院 乳腺・内分泌外科

平 成人

2016/8/20 CSP-HOR年会

RESQ trial

HER2陰性転移・再発乳がん患者を対象にエリブリンとS1の health-related quality of life(HRQoL)を比較するランダム化第Ⅲ相試験

A Phase III, Open Label, Randomized, Controlled, Multicenter Study, comparing Eribulin with S1, to assess Health-related Quality of Life (HRQoL) in the Treatment of Her2 negative Metastatic Breast Cancer Patients

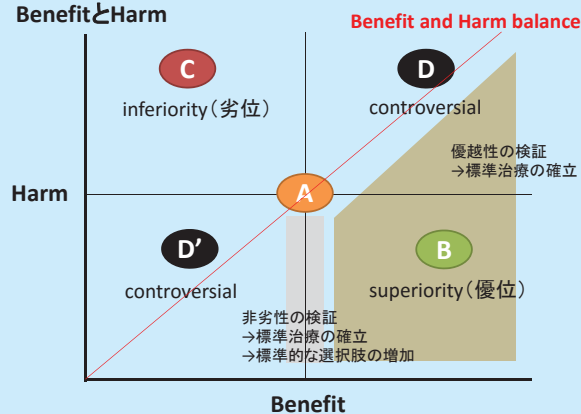
PI 高橋将人先生 北海道がんセンター

✓ 研究の背景-その1

がんの比較臨床研究におけるエンドポイント-HRQoL評価の意義



がんの比較臨床試験におけるエンドポイント BenefitとHarm



がんの治療で評価すべきアウトカムとは？
1995 ASCO ガイドライン

SPECIAL ARTICLE

Outcomes of Cancer Treatment for Technology Assessment and Cancer Treatment Guidelines

Adopted on July 24, 1995 by the American Society of Clinical Oncology*

- 患者アウトカム(生存やQoLなど)を腫瘍関連アウトカム(奏効率など)に優先すべき……
- 複数エンドポイントを考慮すべき。
なぜなら、がん治療の結果を適切に描写する、単一の指標はないから。
- 治療の優劣の議論では、治療による恩恵と、毒性やコストとのバランスが考慮されるべき。

→ 健康関連アウトカムの評価を！！

健康アウトカム(成果、結果、エンドポイント)

◆ Benefit

薬の場合は薬理的有効性から有用性effectiveness
治癒、生存率、生活機能・自立度
健康関連QoL、患者満足度

◆ Harm

有害反応(副作用)、短期・長期(発ガンを含む)
健康関連QoL、患者満足度

限られた資源の中で ……費用対(アウトカム上の)効果の
医療経済評価が求められている

大橋靖雄先生、平成27年CSP-HOR年会資料より

比較臨床研究における
リスク・ベネフィットバランスをどう解釈するか？

the European Society for Medical Oncology
Magnitude of Clinical Benefit Scale (ESMO-MCBS)

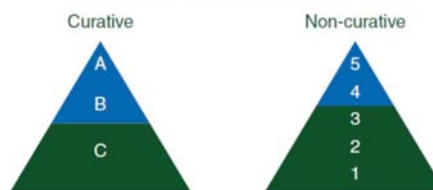
Annals of Oncology 26: 1547-1573, 2015

Table 1. Potential benefits of a new treatment

Living longer	◆ 生存期間とQoL・毒性 ◆ その代替指標として、 -無病生存期間 -増悪生存期間
Improved OS	
Improved surrogate of OS	
DFS (when OS data are immature in adjuvant setting)	
Improved PFS	
Living better	
Improved quality of life	
Improved surrogate of quality of life	
Improved PFS	
Reduced toxicity	

Magnitude of Clinical Benefit Scale

ESMO MCBS evaluation



Curative-Evaluation form 1: for new approaches to adjuvant therapy or new potentially curative therapies

Non-curative-Evaluation forms 2a, b or c: for therapies that are not likely to be curative

ESMO-MCBS grading: OSを主要評価とした試験の解釈

標準治療により、生存期間中央値1年以上の場合

Grade 4	Mark with X if relevant
HR ≤0.70 AND Gain ≥5 months	
Increase in 3 year survival alone ≥10%	

Adjustments

副次評価で、QoLの改善や毒性の軽減が示された場合は、1ランクアップ

Grade 2	
HR >0.70-0.75 QR Gain 1.5-2.9 months	
Increase in 3 year survival alone 3 - <5%	

Grade 1	
HR >0.75 QR Gain <1.5 months	
Increase in 3 year survival alone <3%	

ESMO-MCBS grading: 毒性やHRQoLを主要評価とした場合

Primary outcome is Toxicity or Quality of life AND Non-inferiority Studies

Grade 4	Mark with X if relevant
Reduced toxicity or improved QoL (using validated scale) with evidence for statistical non inferiority or superiority in PFS/OS	

Grade 3	
Improvement in some symptoms (using a validated scale) BUT without evidence of improved overall QoL	

* PFS/OSで非劣性もしくは優越性が示され、毒性の低減、もしくはQoLの改善が認められた場合は、

Grade4: 臨床的な恩恵が大きい、と解釈

✓ 研究の背景-2

なぜ、HRQoLなのか？

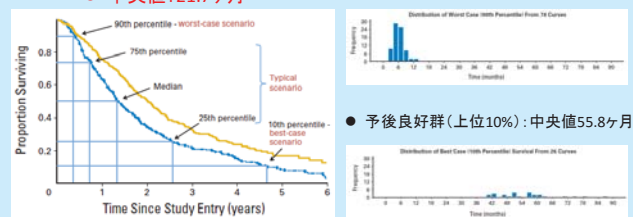


転移乳がんの生命予後: 推定のシナリオ

1999-2009, 一次化学療法に関する36のランダム化比較試験試験毎の生存曲線と全生存期間を統合解析

● 中央値: 21.7ヶ月

● 予後不良群(下位10%): 中央値6.3ヶ月



生命予後が良好な臨床試験の特徴

- HER2陽性で、Trastuzumabの投与を受けてる
- ホルモン受容体陽性が多い

90%以上の対象者が5年以内に死亡

Kiely et al. JCO 2011

がん患者は治療に何を望む？

Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public

Maurice L Slevin, Linda Stubbs, Hilary J Plant, Peter Wilson, Walter M Gregory, P Joanne Armes, Susan M Downer

BMJ VOLUME 300 2 JUNE 1990

- 看護師, 一般人, がん専門医, がん患者を対象とした調査
 - 各々の立場で、化学療法に関するリスク・ベネフィットバランスを調査
- 『どのくらいの効果であれば、あなたは次の治療を受けますか？』

- 副作用の強い治療

悪心、嘔吐、脱毛、倦怠感、月に3-4日の入院

- 副作用の軽い治療

軽度の悪心、脱毛なし、月に1日の通院

このくらいの効果があれば、治療を受ける(中央値)

	看護師 (n=303)	一般人 (n=790)	がん専門医 (n=148)	がん患者 (n=100)
副作用の強い治療(悪心・嘔吐・脱毛・倦怠感, 月に3-4日の入院)を選択				
治療の確率 (%)	50	25	10	1
生命予後改善(月)	24	24	12	12
症状の改善 (%)	50	75	50	10
副作用の弱い治療(軽度の悪心, 脱毛なし, 月に1回の通院)				
治療の確率 (%)	10	10	10	1
生命予後改善(月)	12	12	6	3
症状の改善 (%)	25	25	25	1

Slevin ML et al. BMJ 1990

生命予後を改善しない治療に意味はあるのか？

- Stage I-III期, 術後乳がん患者115人
- ある状況を想定した質問票による調査
- 『再発をし、余命18カ月。どの条件で、どの治療を選択するか』
 - ホルモン療法, 標準的化学療法, 生命に危険が及ぶ実験的治療
 - 予後改善効果は、1週間, 1カ月, 6カ月, 1年, 3年, 5年



『たとえ1カ月でも長生きできるなら、実験的治療を受ける』=15%が選択

『生命予後の改善がなくても、疼痛・症状が軽快する確率が50%なら、標準的化学療法, 実験的治療を受ける』=76%が選択

McQuellon RP et al. JCO 1995

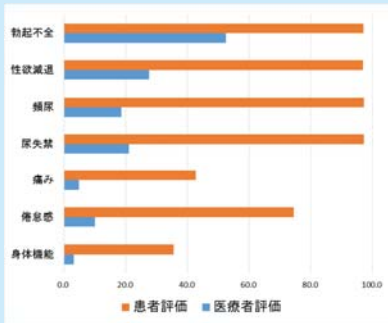
転移乳がん
Goal of therapy

- 生存期間の延長
- 症状のコントロール
- Quality of Lifeの維持改善

これらの評価を
● だれが
● どのように
行うのが妥当か？

患者の症状を、だれが評価すべきか？ 医療者と患者の評価に乖離：その1

前立腺がん患者の症状の評価



% Reporting Impairment Litwin MS et al. J Urol 1998

医療者と患者の評価に乖離：その2

国内、術後補助療法のランダム化比較試験(N-SAS BC02)
タキサンによる化学療法誘発性末梢神経障害の評価の比較

医療者の評価 $\xrightarrow{\text{重度}}$

患者の評価 \downarrow 重度	Kappa=0.16 NCI-CTC Sensory					Kappa=0.02 NCI-CTC Motor				
	0	1	2	3	4	0	1	2	3	4
PNQ Sensory ^a						PNQ Motor ^b				
A	489	38	0	0	0	A	492	5	1	0
B	432	252	4	0	0	B	701	37	2	0
C	113	171	5	0	0	C	231	17	3	0
D	44	66	11	3	1	D	62	10	5	1
E	9	1	0	0	0	E	9	0	0	0

Shimozuma et al. Support Care Cancer 2009 18

患者の「主観」を捉える意義

- 医療者にとって十分には捉えられないが、患者や一般人にとって価値が高い(健康・生活に関する)関心事があり、それを医療に生かす必要があるから。

Shimozuma K, 岡山乳がんチーム医療研究会

✓ 研究の背景-3

転移乳がんの一次化学療法



20

転移乳がん、一次化学療法のこれまで

- 転移乳がんの一次化学療法として、アンスラサイクリン、タキサンが第一選択薬とされてきた。
- 経口5-FU剤、S1とタキサンの比較試験(SELECT-BC)の結果、OSに関してS1の非劣性、QoLに関してS1の優越性が証明され、S1も一次化学療法剤の選択肢となった。

21

Eribulin

- クロイソカイメンから抽出された、ハリコンドリン類の微小管阻害剤
- 転移乳がんの2次以後の治療として、Capecitabine(経口5-FU剤)とEribulinを比較した第Ⅲ相試験(301試験)から、Eribulin群でOSを改善する傾向が認められ、QoLは同等であった。
- Eribulinは一次治療として検証する価値がある。

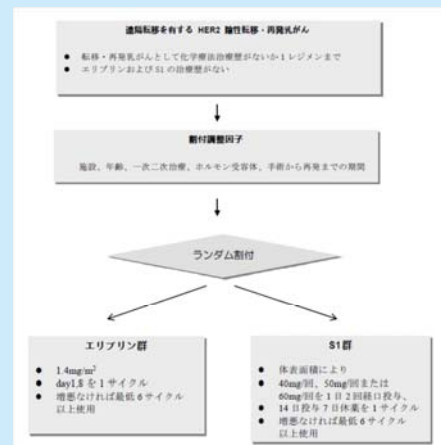
22

✓ 試験デザインの紹介

RESQ trial



23



24

- ✓ RESQ trial, 計画段階で見えてきた課題と対策
 - HRQoLをprimary endpointとする場合-



25

転移乳がんを対象とした比較臨床試験におけるHRQoL評価 システマティックレビュー

- 2001-20014, 進行・転移乳がんを対象とした RCT, n ≥ 50
- HRQoLの評価が行われた試験=49 RCTs
- HRQoLを主要評価とした試験= 5 RCTs
- HRQoL評価の仮説が明記された試験= 10/47 (21%)
- 統計学的事項
 - 統計手法が明記されていない試験= 6/49 (12%)
 - スコアによる有意差検定(時点を問わず)= 26/49 (62%)
 - スコアの臨床的意義を考慮した解釈= 23/49 (49%)

課題: 仮説の明記、臨床的意義を考慮した解析・解釈

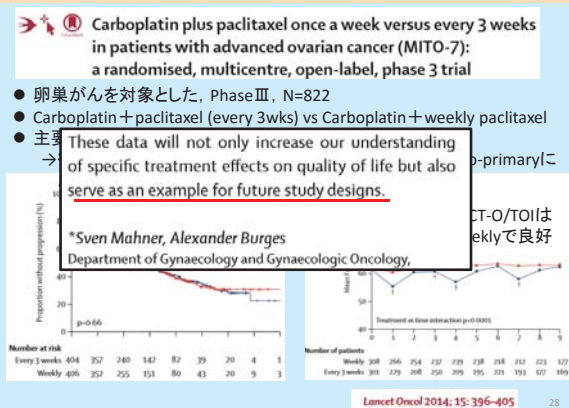
25

QoLを主要評価項目とした比較臨床試験: 乳がん薬物療法

- Leslie GW, et al (BMC cancer, 2011)
 - 術前治療としてweekly Doce.-AC vs tri-weekly Doce.-ACの比較
 - N=89
 - 主要評価項目は, FACT-B TOI
 - 副次評価は, pCR率, 温存率, DFS, OS
 - 結果: 主要評価, 副次評価ともに有意差なし
- Nuzzo F, et al (BMC cancer, 2011)
 - 転移乳がんの一次治療としてweekly Doce. vs tri-weekly Doce.の比較
 - N=139
 - 主要評価項目は, EORTC-QLQ C30のGlobal health status
 - 副次評価は, 奏効率, PFS, OS
 - 結果: Tri-weekly Doce.がQoLにおいて有意に良好。奏効率, PFS, OSは同等。

27

QoLを主要評価項目とした比較臨床試験: 卵巣がん



28

RESQ trial 計画段階での議論から。

- 尺度の選択
- Missing data
- 解析計画

29

課題1: 尺度の選択

- 対照群や治療(薬)の特性, 相違を考慮し, かつ治療選択に重要な意味をもつ(ことが予想される)事象を, もれなく評価できる尺度の選択。
 - **N-SAS BC 05 (AERAS):** SF-36 → 特定の疾患や症状などに特有な健康状態ではなく, 包括的な健康概念を測定
 - **ABROAD:** PNQ (Patient Neurotoxicity Questionnaire) → CIPNの評価指標
- QoLを主要評価項目とする場合
 - 尺度で測定される**複数の指標のうち, どれを優先するか?** (例) FACT-Breastでは, 5つのサブドメインスコア, 3つの要約スコア, 合計8つが算出される。。
 - **多重比較の問題**から, 優先的な指標をしぼる必要がある。良いとこ取りは, 許容されない。
 - 複数エンドポイントの場合は, 解析手順の明確化 検定順序または階層の事前規定, ボンフェローニ法など (FDAガイダンス)

30

RESQ trialでの尺度選択

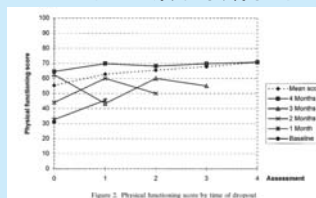
1. EORTC QLQ-C30:がん臨床試験で一般的に用いられるHRQoL尺度(30項目)
2. EORTC QLQ-BR23:乳癌疾患特異的HRQoL尺度(23項目)
3. EQ-5D-5L(EuroQoL 5 Dimension):医療経済評価における効用値測定目的
4. QoLの予期に関する質問: 今後のQoLの予期を問うオリジナル尺度(1項目)
5. 推移の質問: MID(Minimally important difference)の解析に必要な主観的アンカー情報として用いる。
6. 患者選好: 治療の有害反応に関して, 患者の選好を問うオリジナル尺度

Primary endpointはEORTC QLQ-C30のGlobal Health Statusに設定

31

課題2: Missing data, 欠測

- QoLデータの回収, コンプライアンスは最重要課題。
- Missing data
 - **item non-response:** 調査票のうち, いくつかの項目が欠測
 - **unit non-response:** 調査票が丸々欠測
 - intermittent missing, dropout missing, late entry
 - 予後不良の疾患の場合には, 状態の悪化や死亡による欠測
- なぜMissing dataが重要視されるのか?
 - バイアスの原因となり得るから



- EORTC QLQ-C30, 毎月, 5回の調査
- 全体の平均スコアは増加傾向
- しかし, drop out(月)で層別化した平均スコアは, drop outが早いほど, 低い!!

Statist. Med., 17, 697-709 (1998)

32

Missing data: 欠測のつづき

- Missing dataのメカニズム
 - Missing Completely At Random (MCAR): 完全にランダムな脱落
 - Missing At Random (MAR): 脱落確率が観測データにのみ影響を受ける
 - **Missing Not At Random (MNAR): 脱落確率が欠測データに影響を受ける**
- Missing dataの要因
 - **最大要因は施設**
 - スイス、肺がんを対象としたRCTで、QoL調査票未回収に及ぼした唯一の要因は、参加施設のみであった (Hürny C, et al. Ann Oncol. 1992)
 - 病状の増悪
 - 英国、肺がんを対象としたRCTで、病状の進行とともに、QoL調査票への記入は困難になったが、短い項目では継続可能 (Geddss, et al. Eur J Cancer 1990)
 - 患者の死亡
 - 死亡による欠測をどう扱うかは、定まっていない(はず)。
- 対策
 - 無理のない調査計画、欠測の扱いを明記、解析での工夫、**理由の調査**

33

RESQ trial, Missing dataへの対策

- リマインドと欠測理由の調査
 - 1) 対象者の病状の悪化
 - 2) 対象者の死亡
 - 3) 対象者の調査拒否
 - 4) 対象者の記入・送付忘れ・紛失
 - 5) 対象者の要因による、上記以外の理由
 - 6) 担当者から対象者への調査票の配布忘れ
- 病状の増悪、死亡による欠測を「臨床的に意味のあるQoLの低下」と定義、primary endpointのイベントとしてカウント

34

課題3: QoLスコアの解釈

臨床的に意味のある差とは？

- 尺度スコアの統計学的有意差 \neq 臨床的に意味のある差
- 結果の解釈には、何点の差が臨床的に意味のある差か (= 明らかな改善、増悪) を定義する必要がある。
 - 例: 7-point scale で2点上がれば確実に改善したと言える

→ Responder?

-3 -2 -1 ±0 +1 +2 +3



+2以上を改善と定義することは妥当か？

35

MIDの定義

- PRO尺度におけるMinimally Important Difference (MID)の理解 = 臨床的に意味のある最小の差をどのように定義するか

Anchor based approach

臨床的に意味のある事象を基準にMIDを同定する方法
Anchorの例: PS, 臨床病期, 推移の質問による主観的な評価

- EORTC QLQ-C30のMID (100点換算)

- 5-10ポイント わずかな差
- 10-20ポイント 中等度の差
- 20ポイント以上 重大な差

(Osoba D, et al. JCO 1998)

RESQ trialでの解析計画

- 2群間の比較には minimally important difference (MID) を加味した解析を実施。EORTC-QLQ C30 の Global Health Status の MID を 10 ポイントと規定
- 各調査時点で、EORTC-QLQ C30 Global Health Status のスコアが、ベースラインより10ポイント以上の低下を認めた場合は「臨床的に意味のあるQoLの低下イベント: QoL 低下例」と定義
- 死亡に伴う欠測 および病状の悪化に伴う欠測は「臨床的に意味のあるQoLの低下イベント: QoL 低下例」として扱う。
- 群間比較には、登録後1年以内に発生したQoL低下例の割合の差を用いる。各群の登録後1年までの累積QoL低下割合をKaplan-Meier法により求め、群間差および信頼区間を求める。

37

結語

HRQoLをprimary endpointとした新規臨床試験、RESQ trialを紹介し、その計画段階における課題と対策を報告した。

38