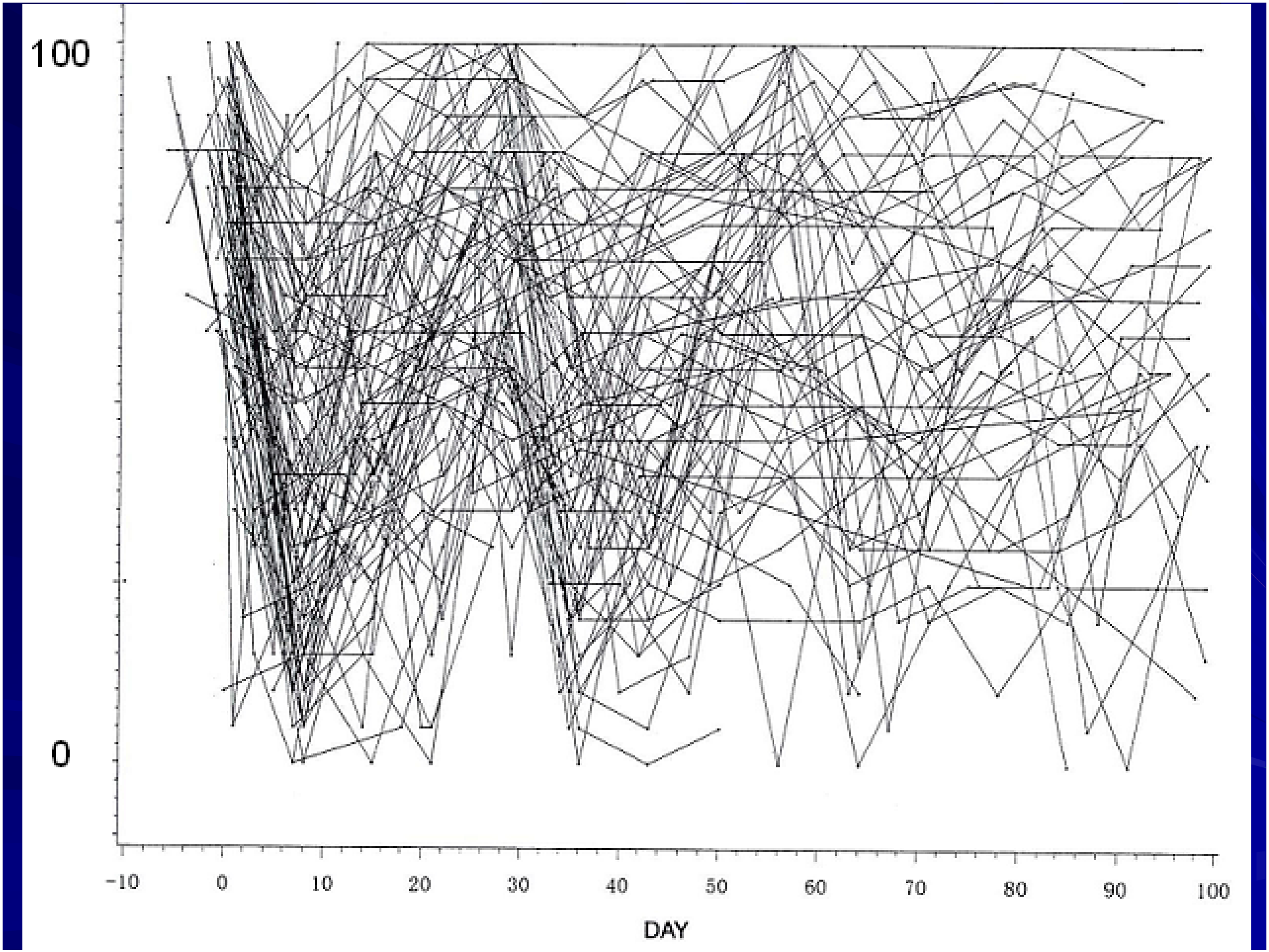
Health-Related Quality of Life（HRQOL）の評価

森田 智視
横浜市立大学
大学院医学研究科／附属市民総合医療センター
臨床統計学・疫学
E-mail: smorita@urahp.yokohama-cu.ac.jp

1


100
0
−10
0
10
20
30
40
50
60
70
80
90
100
DAY

QOLは結果変数の一つ

要因

- バクロ
  - 大気汚染
  - 喫煙
  - 電磁波
- 治療
  - 薬
  - 手術

観察

結果

- 喘息
- 肺癌
- 小児白血病

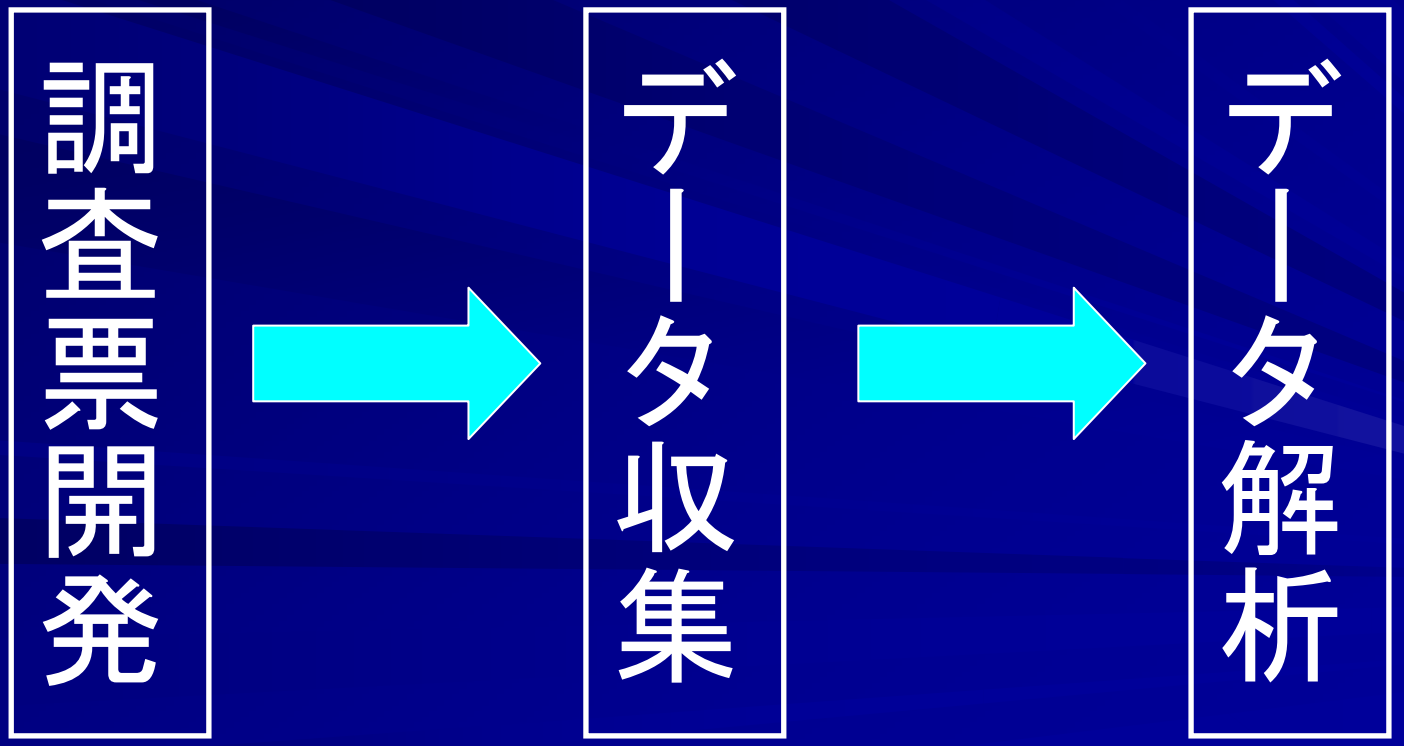
- 疾病治癒
- 死亡（延命）
- 副作用
- **QOL**

3

本日の目標

QOL調査・研究の全体像を理解する

-統計学的側面を中心に-



4

内容

1. QOLの定義と評価:「QOLって何？」
   - QOLの定義
   - QOL調査票の開発
2. QOLデータ解析:「どう測定・解析するの？」
   - QOLの測定
   - データ解析方法

5

QOLの定義

なぜQOLに注目？

- がんや糖尿病などの慢性疾患
  - 複数の競合する治療選択肢があることが多い
  - がん：完全な治癒が望めない疾患
  - 糖尿病：ライフスタイル変容が治療の一環として求められる
  - 治療法の選択がその後の生活の質に大きな影響
- 客観的エビデンスだけでは十分ではない
  - 有効性（例：生命予後）や安全性（例：副作用）
- 医師だけでなく患者も納得できる治療選択
  - 患者個人の価値観に基づく評価指標を用いた治療法評価

7

では、どうすればよい？

- 患者自身から発信される主観的な情報を定量化しつつ
- データを収集し
- 適切な統計手法により解析し、適切に提示することが必要

言うのは簡単！

8

QOL研究のむずかしい点

- 実体のないものを測定する
  - 白血球数なら採血して単位量あたりの数
  - 生存時間なら死亡まで追跡調査すればよい
  - Quality of life？
    - 何となくわかるが、いったい何もの？
- 苦労の多いデータ回収
  - ただでさえ大変なデータ回収
  - 調査票に記入してもらい、回収しなければならない
- 厄介なデータ解析
  - 多ドメイン・多時点データ
  - データ欠測：　健康状態悪化して記入できない

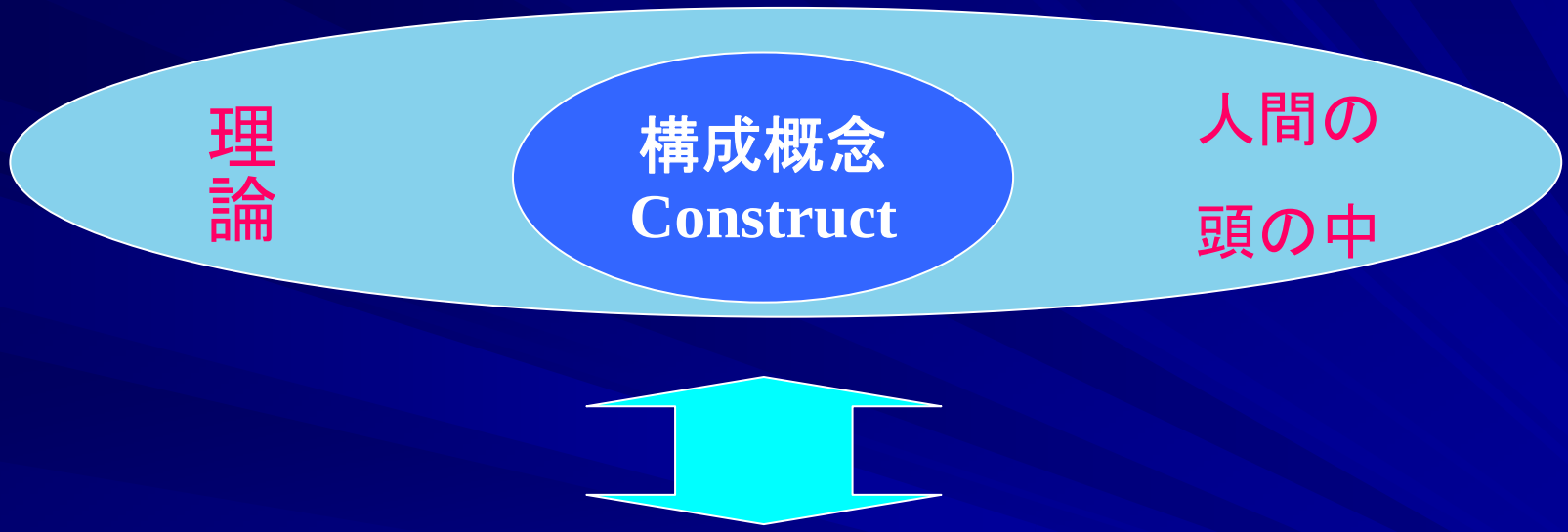
9

そもそもQOLとは？

- QOLは極めて抽象的な概念
  - 第三者評価による症状や日常活動度を意味する場面にまで気軽に使用
  - QOLの定義と測定方法について1980年代より活発な議論（定義についての表現は様々）
- コンセンサス
  - 抽象的な構成概念（construct）とし、活動性、身体性、精神性、社会性などの多次元構造
  - 評価は患者本人による主観的なもの

10

構成概念?

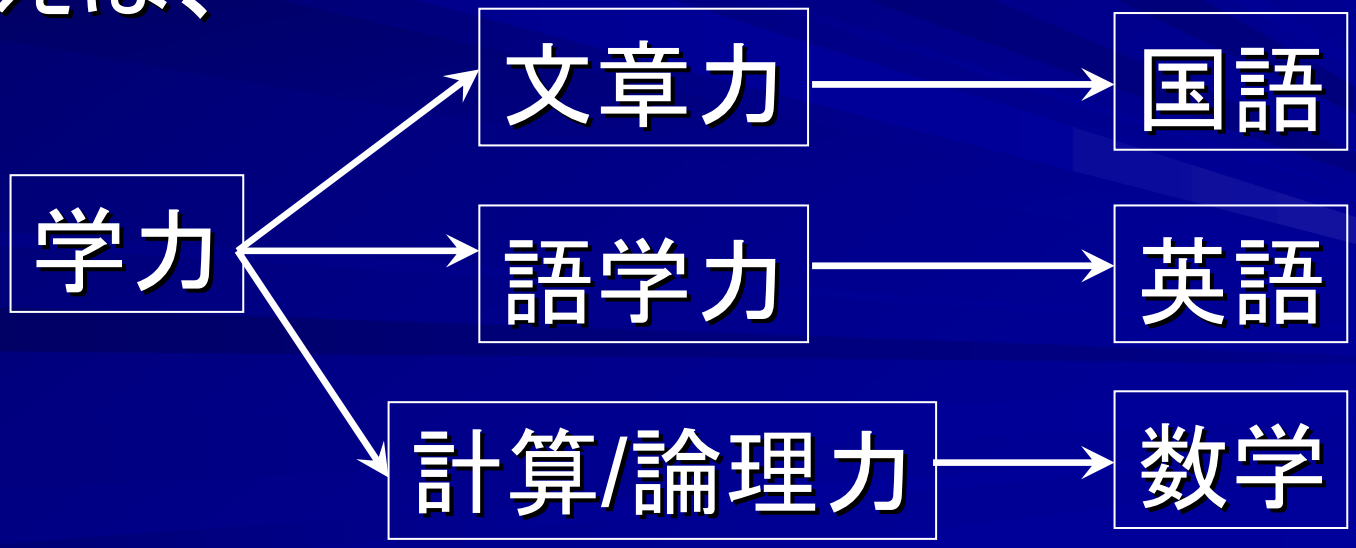


11

構成概念→観察

- 上位概念に定義を与える
- 概念を操作化（operationalize）し、実際に測定可能なものに変化させる

例えば、



12

健康関連（Health-Related）QOL

QOLそのものはとても大きな概念
　人生観や宗教的要素も含まれる

Health-related：疾病や医療介入によって影響を受ける部分

(HR)QOLの階層性と多次元性

QOL調査票

どのようにして測る？
-自記式調査票（尺度）の開発-

- 包括的（generic）な尺度
  - 少なくとも身体性、活動性、精神性、社会性の4要素から構成
  - 複数ドメインに対応する複数項目
  - EORTC-QLQC30、FACT-G、SF36

- 疾患特異的な尺度（モジュール）
  - 対象疾患による症状あるいは患者が感じる「問題」に焦点

例：がん領域

  - 肺癌：EORTC QLQ-LC13、FACT-L
  - 乳癌：EORTC QLQ-BR23、FACT-B
  - 胃癌：EORTC QLQ-STO22、FACT-GC
  - Taxane系薬剤の毒性評価：FACT-TAXANE　など

16

がん領域：欧米では

1980年代より欧米において盛んに

- FACT-G
  - 身体面、社会面、精神面、活動面の4ドメイン
- EORTC QLQ-C30
  - 身体面、役割面、認知面、心理面、社会面の5ドメイン
  - 数個の症状に関するスケール（疲労、疼痛、悪心・嘔吐など）
  - 総括的な（グローバル）QOL

17

日本では、

日本における開発は1980年代後半から

2つのアプローチ

- 欧米開発QOL調査票の翻訳
  EORTCやFACTの日本語版

- 日本独自のQOL調査票開発
  栗原班QOL調査票（QOL-ACD）
  活動性、身体性、精神性、社会性の4ドメイン
  グローバルQOL

18

癌種特異的モジュール -EORTC-

http://www.eortc.be/home/qol/

Validated modules:

- Lung (LC13)
- Breast (BR23)
- Head & Neck (H&N35)
- Oesophageal (OES18)
- Ovarian (OV28)

Modules in phase IV:

- Prostate
- Gastric
- Multiple Myeloma
- Satisfaction with Care

Modules in Phase III:

- High Dose Chemotherapy
- Chronic Lymphocytic Leukemia
- Superficial Bladder
- Muscle Invasive Bladder
- Brain
- Colorectal
- Ophthalmic
- Pancreatic

19

癌種特異的モジュール -FACT-

http://www.facit.org/

- FACT-B: For patients with Breast cancer
- FACT-Bl: For patients with Bladder cancer
- FACT-Br: For patients with Brain cancer
- FACT-C: For patients with Colorectal cancer
- FACT-CNS: For patients with cancer in the Central Nervous System
- FACT-Cx: For patients with cancer of the Cervix
- FACT-E: For patients with Esophageal cancer
- FACT-En: For patients with Endometrial cancer
- FACT-H&N: For patients with Head & Neck cancer
- FACT-Hep: For patients with Hepatobiliary cancer (liver, bile duct and pancreas)
- FACT-L: For patients with Lung cancer
- FACT-Leu: For patients with Leukemia
- FACT-Lym: For patients with Lymphoma (NHL)
- FACT-O: For patients with Ovarian cancer
- FACT-P: For patients with Prostate cancer
- FACT&GOG-Ntx: For patients with Neurotoxicity
- FACT-BMT: For patients undergoing Bone Marrow Transplantation
- FACT-BRM: For patients receiving Biologic Response Modifiers
- FACT-Taxane: For patients receiving Taxane therapy

20

どのようなものか実際に見てみましょう
FACT-B

FACT-B（第 4-A 版）

下記はあなたと同じ症状の方々が重要だと述べた項目です。項目ごとに、ごく最近（過去７日間程度）のあなたの症状に最もよくあてはまる番号をひとつだけ選び○で囲んでください。

	身体症状について	全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
GP1	体に力が入らない感じがする。	0	1	2	3	4
GP2	吐き気がする。	0	1	2	3	4
GP3	体の具合のせいで家族への負担となっている。	0	1	2	3	4
GP4	痛みがある。	0	1	2	3	4
GP5	治療による副作用に悩んでいる。	0	1	2	3	4
GP6	自分は病気だと感じる。	0	1	2	3	4
GP7	体の具合のせいで、床（とこ）（ベッド）で休まざるを得ない。	0	1	2	3	4

	社会的・家族との関係について	全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる

GP7	体の具合のせいで、床(とこ)(ベッド)で休まざるを得ない。	0	1	2	3	4

社会的・家族との関係について

		全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
GS1	友人たちを身近に感じる。	0	1	2	3	4
GSX1	家族を親密に感じる。	0	1	2	3	4
GS2	家族から精神的な助けがある。	0	1	2	3	4
GS3	友人からの助けがある。	0	1	2	3	4
GS4	家族は私の病気を充分受け入れている。	0	1	2	3	4
GS5	私の病気について家族間の話し合いに満足している。	0	1	2	3	4
GSX2	私は病気ではあるが、家族の生活は順調である。	0	1	2	3	4
GS6	パートナー（または自分を一番支えてくれる人）を親密に感じる。	0	1	2	3	4
Q1	次の設問の内容は、現在あなたの性生活がどの程度あるのかとは無関係です。答えにくいと思われる場合は四角に ✓印を付け、次のページの設問に進んで下さい。					
GS7	性生活に満足している。	0	1	2	3	4

Japanese　　10/24/1997

© 1987, 1997

FACT-B (第 4-A 版)

項目ごとに、ごく最近（過去７日間程度）のあなたの状態に最もよくあてはまる番号をひとつだけ選び、○で囲んでください。

精神的状態について

		全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
GE1	悲しいと感じる。	0	1	2	3	4
GE2	病気を冷静に受け止めている自分に満足している。	0	1	2	3	4
GE3	病気と闘うことに希望を失いつつある。	0	1	2	3	4
GE4	神経質になっている。	0	1	2	3	4
GE5	死ぬことを心配している。	0	1	2	3	4
GE6	病気の悪化を心配している。	0	1	2	3	4

活動状況について

		全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
GF1	仕事（家のことも含む）をすることができる。	0	1	2	3	4
GF2	仕事（家のことも含む）は生活の張りになる。	0	1	2	3	4

QE3	病気と闘うことに希望を失いつつある。.....	0	1	2	3	4
QE4	神経質になっている。.....	0	1	2	3	4
QE5	死ぬことを心配している。.....	0	1	2	3	4
QE6	病気の悪化を心配している。.....	0	1	2	3	4

活動状況について

		全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
QF1	仕事（家のことも含む）をすることができる。	0	1	2	3	4
QF2	仕事（家のことも含む）は生活の張りになる。..	0	1	2	3	4
QF3	生活を楽しむことができる。.....	0	1	2	3	4
QF4	自分の病気を充分受け入れている。.....	0	1	2	3	4
QF5	よく眠れる。.....	0	1	2	3	4
QF6	いつもの娯楽（余暇）を楽しんでいる。.....	0	1	2	3	4
QF7	現在の生活の質に満足している。.....	0	1	2	3	4

FACT-B (第 4-A 版)

項目ごとに、ごく最近（過去７日間程度）のあなたの状態に最もよくあてはまる番号をひとつだけ選び、○で囲んでください。

その他心配な点

		全くあてはまらない	わずかにあてはまる	多少あてはまる	かなりあてはまる	非常によくあてはまる
B1	息切れがする。・・・・・・・・・・・・・・・・・	0	1	2	3	4
B2	どのような服をどう着るか、人目が気になる。・	0	1	2	3	4
B3	片方あるいは両方の腕に腫れまたは痛みがある。	0	1	2	3	4
B4	女性としての魅力があると思う。・・・・・・・・	0	1	2	3	4
B5	脱毛に悩まされている。・・・・・・・・・・・・	0	1	2	3	4
B6	家族の他の者がガンにかかるのではないかと心配である。・・・・・・・・・・・・・・・・・・・・・	0	1	2	3	4
B7	ストレスがたまると病気に影響が出るのではないかと心配である。・・・・・・・・・・・・・・・・	0	1	2	3	4
B8	体重の変化に悩んでいる。・・・・・・・・・・・	0	1	2	3	4
B9	女であると感じることができる。・・・・・・・・	0	1	2	3	4

まず、QOL調査票で用いられる主なスケールを整理してみましょう！

- VAS（Visual Analog Scale）
- リッカートスケール（Likert scale）
- フェイス・スケール（Face scale）

26

VAS
(Visual Analog scale)

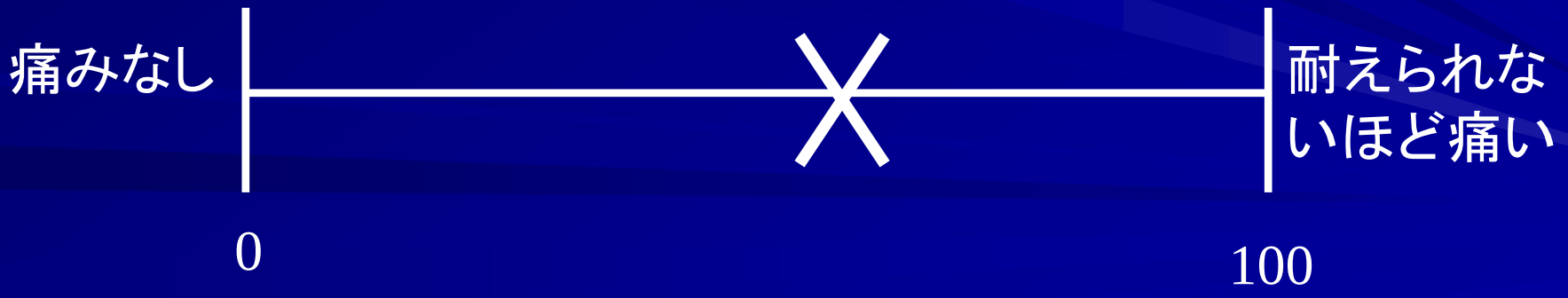
- 通常、100mmの長さの線を使用
- 両端に、極端な項目(アンカー)を置く
- 線上に縦線を引く、またはチェックを付けるなどして、回答者の状態を表してもらう
- 端からの長さを測定して得点化する。

27

VAS
(Visual Analog scale)

例：あなたの頭痛についておうかがいします。

現在の痛みはどの程度ですか。下線に印を付けて表して下さい。



28

リッカートスケール（Likert scale）

- 最も代表的な連続判定尺度
- 通常、５ or ７段階の選択肢
  - どれくらいあてはまるか
  - どれくらい賛成か
  - どれくらいの頻度で経験するか、など
- 各項目の回答に点数が与えられ、その合計点を回答者の総合点（ドメインスコア）とする

29

リッカートスケール（例）

SF36から抜粋

次にあげた各項目はどのくらいあなたにあてはまりますか。

	まったくそのとおり	ほぼあてはまる	何とも言えない	ほとんどあてはまらない	ぜんぜんあてはまらない
1) 私は他の人に比べて病気になりやすいと思う	1	2	3	4	5
2) 私は、人並みに健康である	1	2	3	4	5
3) 私の健康は、悪くなるような気がする	1	2	3	4	5
4) 私の健康状態は非常に良い	1	2	3	4	5

30

リッカートスケール：Scoring

1) 私は他の人に比べて病気になりやすいと思う	1	**2**	3	4	5
2) 私は、人並みに健康である	1	2	3	**4**	5
3) 私の健康は、悪くなるような気がする	**1**	2	3	4	5
4) 私の健康状態は非常に良い	1	2	3	4	**5**

小 ← スコア → 大
悪 ← QOL → 良 になるように変換

Scoring：　2 + 2 + 1 + 1 = 6点

31

※参考　得点の解釈:0-100得点化

例:5段階選択肢の4項目:ドメインスコアの計算

raw Score:　2 + 2 + 1 + 1 = 6点

4項目全て最小値なら
↓
1+1+1+1=4

全て最大値なら
↓
5+5+5+5=20

(0-100得点化 ： (6-4)/ (20-4)×100=12.5点)

32

フェイス・スケール

ここ数日間の状態に相当する顔の番号に○をつけて下さい。

33

QOL調査票の開発

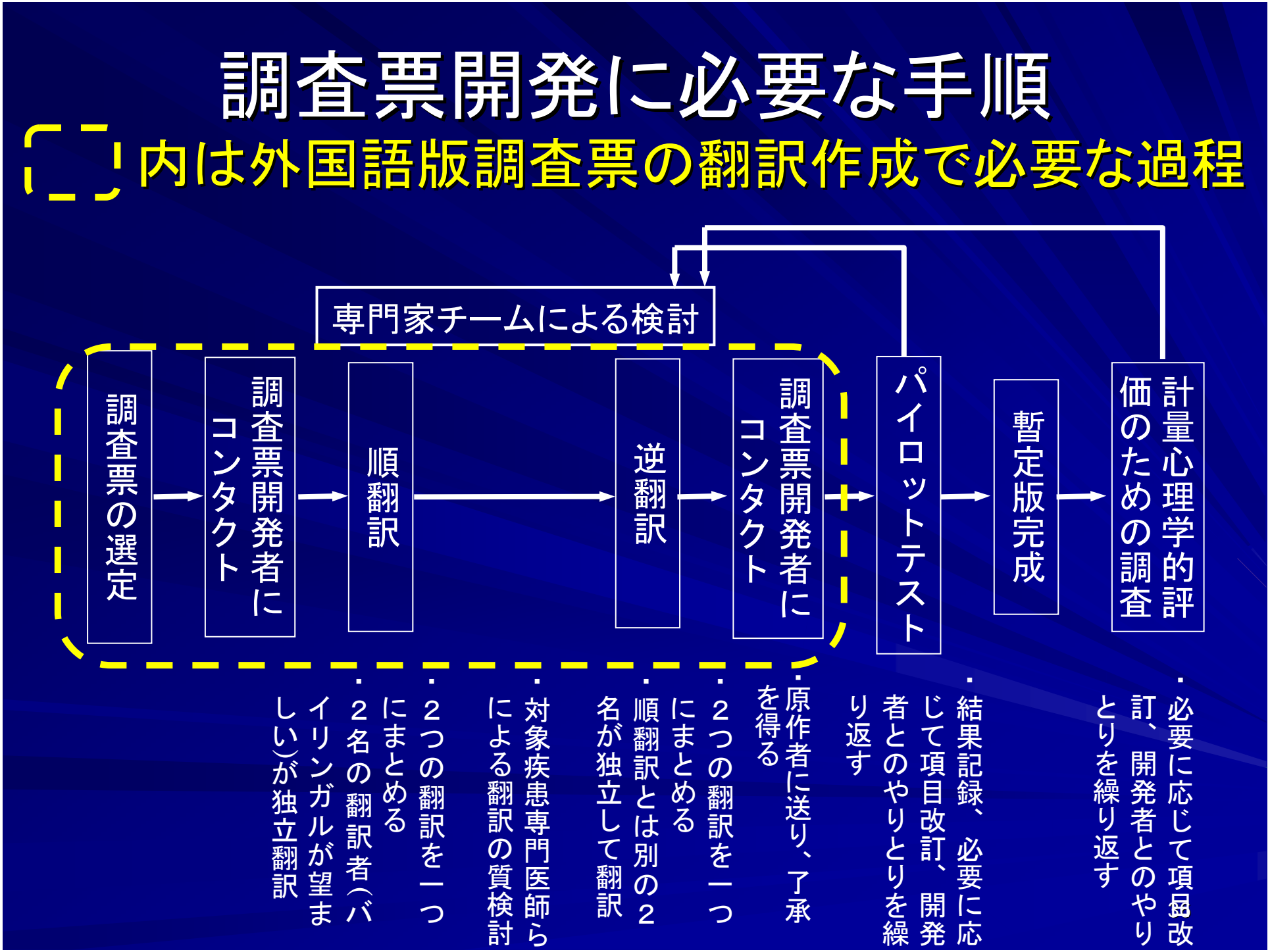
調査票はどうようにして開発？

- 海外で開発された調査票の翻訳
  - 日本語版の作成
  - EORTC、FACTの日本語版
- 新しく開発

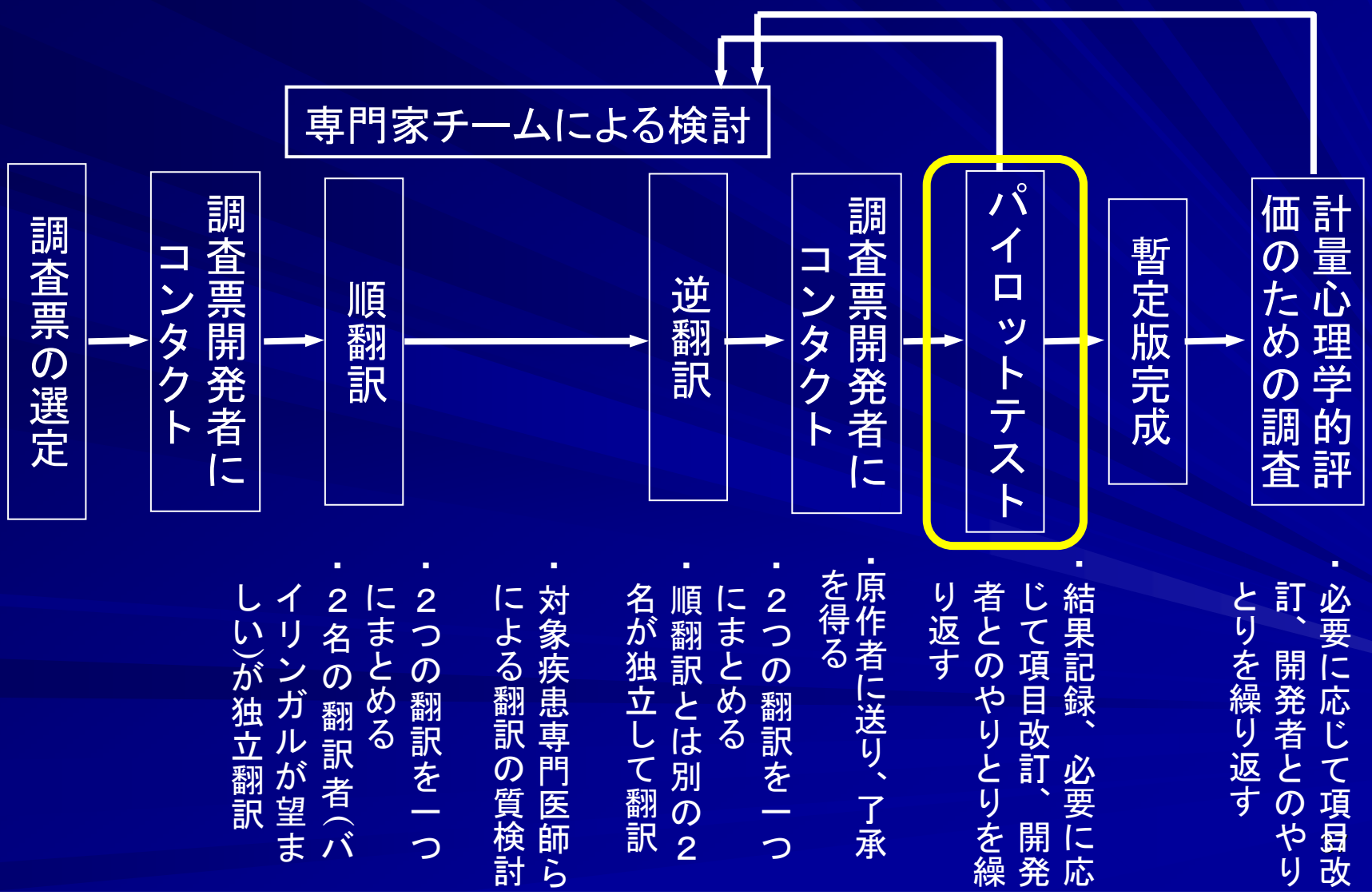
35

調査票開発に必要な手順

内は外国語版調査票の翻訳作成で必要な過程



調査票開発に必要な手順



パイロットテストの目的

- 内容的妥当性の検討
  - 意味がよく伝わらない項目がないか
  - 内容が網羅されているか
- 実施可能性の評価
  - 調査実施上の問題点はないか

必要に応じて項目改訂
開発者とのやりとり

パイロットテストの実施手順

- 対象患者5～10人に実際に調査票記入を依頼
  - 可能な限り、重症度などがバラツクように
- 回答後にインタビューあるいはコメント記入
  1) 回答に戸惑う部分
  2) 答えにくい部分
  3) 分かりにくい部分
- 改善案の提案：　多くは文化差の調整

39

パイロットテストの事例
胃癌用モジュール：EORTC QLQ-STO22日本語版

- 胃癌患者10人
  - 男6人、平均年齢56（23-73）、PS 0-4、病期I～IV
  - 記入時間：20～40分
- インタビュー結果を（改善案とともに）まとめたレポートを提出：
  EORTCのQOL Unitと検討・協議
- 質問文の修正→STO22日本語版の完成

40

レポート

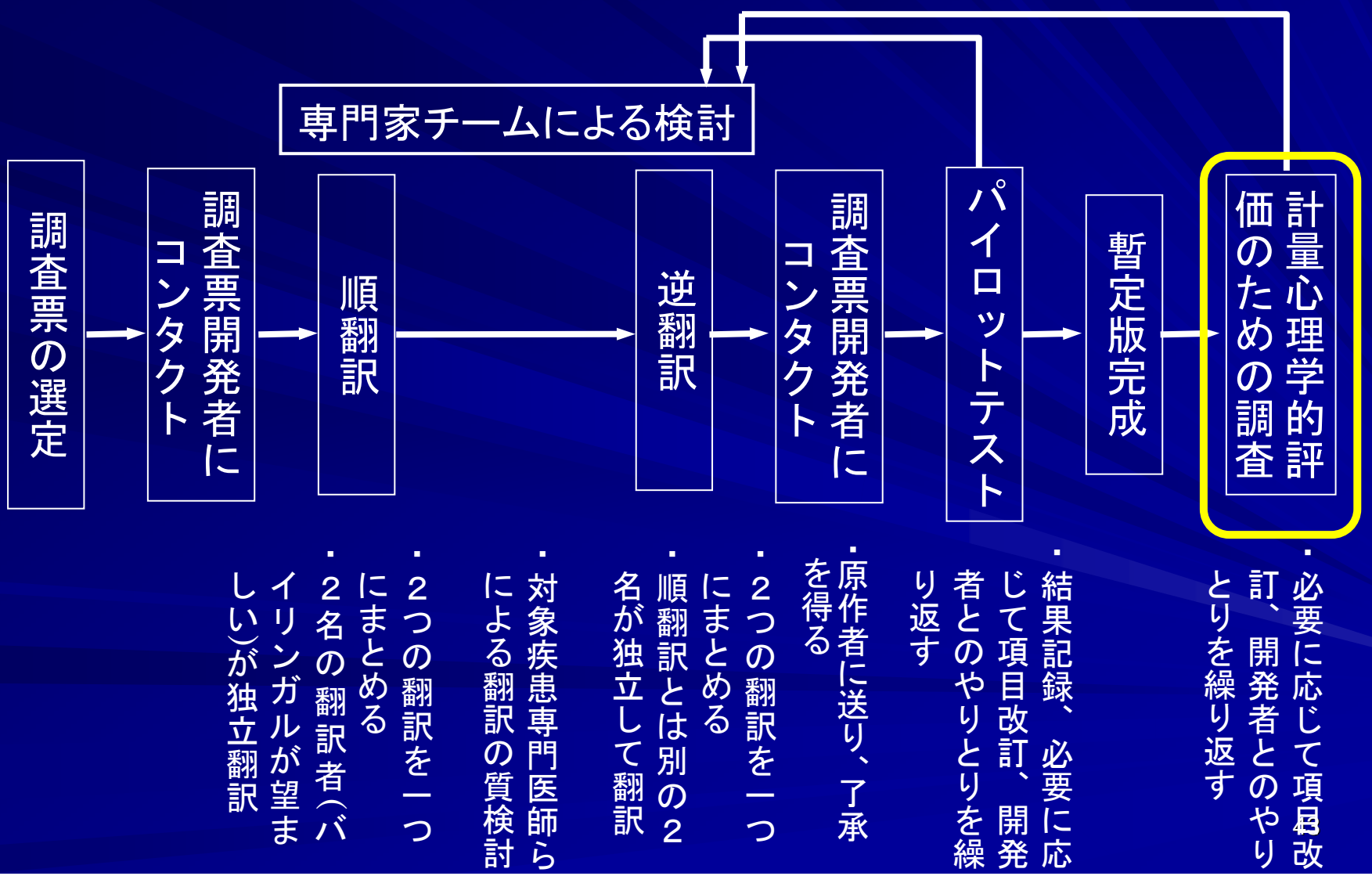
- 全体的に理解しやすい質問文で構成されている
- 記入時間も負担を感じるほど長いものではない
- 22項目の質問内容は、胃癌患者の健康状態を適切かつ包括的に表わしている
- 質問内容明確化のため質問文を修正（5項目）
  - 項目31：“固形食”の例として“ごはんやおかずなど”を追記
  - 項目32：“流動食”の後に例として“おかゆなど”を追記
  - 項目35：“胃のあたり”は、胃の切除を行った患者にとっては理解しにくい→“おなか”を追記
  - 項目37：“腹部に膨張感”は意味が不明瞭→“腹部に張った（ガスが溜まっているような）かんじ”に修正
  - 項目42：“食事を普通に楽しめない”の“普通に”は意味が不明確→め“いつものように”に修正

41

修正を加えた5項目

項目番号	修正前	修正後
31	固形食を食べるのに問題がありましたか。	固形食(ご飯やおかずなど)を食べるのに問題がありましたか。
32	流動食ややわらかいものを食べるのに問題がありましたか。	流動食(おかゆなど)ややわらかいものを食べるのに問題がありましたか。
35	胃のあたりに痛みを感じましたか。	胃のあたり(おなか)に痛みを感じましたか。
37	腹部に膨張感がありましたか。	腹部に張った(ガスが溜まっているような)かんじがありましたか。
42	食事を普通に楽しめないことがありましたか。	食事をいつものように楽しめないことがありましたか。

調査票開発に必要な手順



調査票バリデーション解析の“お作法”

- 計量心理学的解析
  - 実施可能性/項目分析　feasibility/item analysis
  - 信頼性　reliability
  - 妥当性　validity
  - 感度/反応性　sensitivity/responsibility

お作法 その1

- 計量心理学的解析
  - 実施可能性/項目分析　feasibility/item analysis
  - 信頼性　reliability
  - 妥当性　validity
  - 感度/反応性　sensitivity/responsibility

実施可能性/項目分析

- 未記入項目数の検討
  - 欠測割合の計算
- 記入に要する時間（回答者の負担）
  - 平均、Min、Max
- 回答の偏り（度数分布）
  - ヒストグラム（各選択肢への回答割合）

46

実施可能性/項目分析

未記入項目数の検討

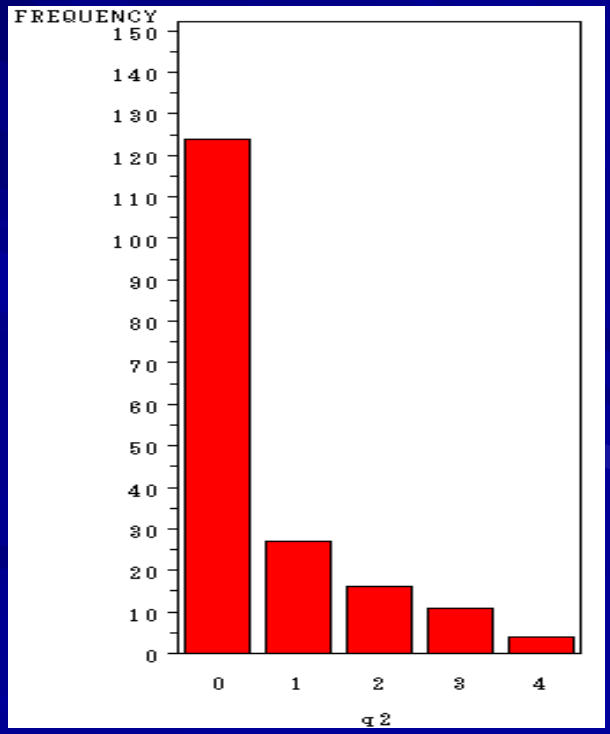
- 栗原班（項目6）：肺癌
  - ひとりで風呂に入ることができましたか？
    - 他に比べ未回答が多かった

- FACT-G(S7)
  - 性生活に満足している
    - 回答割合：50%～60%程度

47

実施可能性/項目分析

- 記入に要する時間(回答者の負担)
  - 対象疾患を考慮
    - 例:進行がん
      長くても30分以内
      15分程度が理想
- 回答の偏り(度数分布)
  - 術後乳癌:FACT-G(P2)
    - 吐き気がする



全くあてはまらない　　非常にあてはまる

実施可能性/項目分析

不適切だと判断された場合

項目の改良あるいは除外

49

信頼性／妥当性を調べる3ステップ

1. ドメイン構成の設定

   どの項目がどのドメインに属するか？

2. 調査票内部構造を検討

   測定データを解析し、・項目とドメイン、・ドメインとドメインの関係を調べる

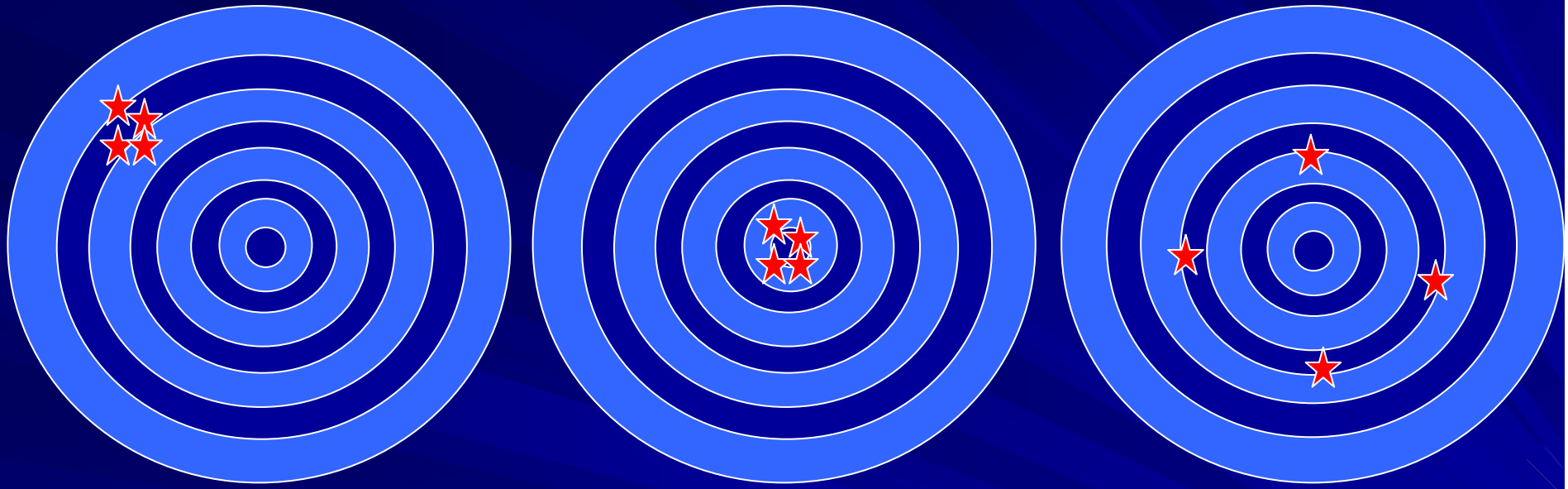
3. 外部基準との関係を検討

   ドメインと調査票外データとの関係を調べる

50

測定の信頼性と妥当性

QOLに特化した話しではなくデータ測定の一般的話題です



信頼性は高い
妥当性は低い

信頼性も高い
妥当性も高い

信頼性は低い
妥当性は高い

信頼性:「同じモノを測っているか？」
妥当性:「測りたいモノを測っているか？」

51

お作法 その2

計量心理学的解析

- 実施可能性/項目分析　feasibility/item analysis
- 信頼性　reliability
- 妥当性　validity
- 感度/反応性　sensitivity/responsibility

52

信頼性

「同じモノを測っているか？」

test-retest 再現性（reproducibility）

- 対象者の状態が変化していないと想定される期間に（通常は）2回測定
- 結果が一致する（再現される）かどうか
  - 繰り返し評価の適切な間隔は、対象疾患（病態進行速度）と治療変更などを考慮
    - 進行がん：2～7日間隔
    - 糖尿病や軽症の喘息：2週間-1月間隔

κ（カッパ）係数、級内相関係数を計算

κ（カッパ）係数

順序尺度変数を含むカテゴリカル変数の一致の度合い（偶然の一致の影響を調整）

1：完全一致、0：偶然一致と同程度（負の値も取りうる）

0未満：poor
0～0.20：slight
0.21～0.40：fair
0.41～0.60：moderate
0.61～0.80：substantial
0.81～1：almost perfect

(参考文献：Landis et al., 1977)

54

級内相関係数
intra-class correlation coefficient

連続変数の再現性評価に用いられる標準的な指標

QOLスコアが、真値と誤差の和で与えられると想定

誤差→0　級内相関係数→1（最大値）

級内相関係数が1に近いほど再現性（精度）↑

55

栗原班調査票各ドメインの test-retest再現性 (肺癌 n=107)

測定時期および間隔
治療開始前の2～7日間

ドメイン	項目数	級内相関係数
活動性	5	0.61
身体性	5	0.71
精神性	4	0.79
社会性	5	0.87

お作法 その3

計量心理学的解析

- 実施可能性/項目分析　feasibility/item analysis
- 信頼性　reliability
- 妥当性　validity
- 感度/反応性　sensitivity/responsibility

57

妥当性

「測りたいモノを測っているか？」

- 内容的妥当性　content (face) validity
- 構成概念妥当性　construct validity
  - 収束的妥当性　convergent validity
  - 弁別的妥当性　discriminant validity
  - 因子妥当性　factor validity
  - 内的整合性　internal consistency
- 基準関連妥当性　criterion validity
  - 予測妥当性　predictive validity
  - 臨床的妥当性　clinical validity

58

内容的妥当性

測定したいと思っている内容と(を)

作成した項目

■合致しているか？

- 測定したい内容から離れた項目はないか?
- 内容が明確か?
- 文章の意味がうまく伝わるか?

■網羅的に捉えているか？

- 患者の問題を網羅しているか?
- 過剰な重複がないか?

専門家チームによる検討、小規模な調査、他医療スタッフの意見聴取
いわゆる「質的研究」の範疇：　統計処理は使われない

59

構成概念妥当性

「構成概念」と測定データとの関係を調べる

あらかじめ理論的に定めた構成概念「ドメイン構造」	⟺	実際に測定された結果（データ）

60

構成概念妥当性
-収束的／弁別的妥当性-

収束的妥当性

各項目が帰属“する”と想定されたドメインと相関が十分“高い”か?

弁別的妥当性

各項目が帰属“しない”と想定されたドメインと相関が十分“低い”か?

Multitrait scaling analysisを用いるのが一般的

各項目得点が帰属ドメイン得点に重複して含まれないように当該項目を除いた上で相関係数行列を算出する

61

構成概念妥当性
-Multitrait scaling analysis-

強い理論的根拠はないが、

収束妥当性

帰属するドメインとの相関が0.4以上であること

弁別妥当性

帰属ドメインより他のドメインとの相関が高い場合エラーと判断

他ドメインとの相関＞帰属ドメインとの相関係数＋標準誤差×2

62

収束的妥当性・弁別的妥当性の例
栗原班調査票（進行肺癌）

ドメイン	項目数	相関係数の範囲 収束的妥当性	相関係数の範囲 弁別的妥当性	収束的妥当性テスト 基準を満たした項目数／全数	収束的妥当性テスト %	弁別的妥当性テスト 基準を満たした項目数／全数	弁別的妥当性テスト %
活動性	6	.64-.80	.26-.58	6/6	100	18/18	100
身体性	5	.37-.74	.10-.62	4/5	80	15/15	100
精神性	5	.35-.67	.18-.54	4/5	80	15/15	100
社会性	5	.45-.75	.15-.48	5/5	100	15/15	100

63

ドメイン間の弁別性
Distinctiveness among the domains

ドメイン間相関係数（Inter-scale correlation）が高すぎない（0.7以下）ことを示す

研究者の頭の中では“別もの”でも、実は同じようなものであったりしないか？

参考文献：Aaronson NK et al. (1993) J Natl Cancer Inst, Kaasa S et al. (1995) Eur J Cancer

ドメイン間のピアソン相関係数　（栗原班調査票：進行肺癌）

ドメイン	活動性	身体性	精神性	社会性
活動性	1	0.59	0.62	0.33
身体性	-	1	0.59	0.27
精神性	-	-	1	0.52
社会性	-	-	-	1

64

構成概念妥当性
-因子妥当性-

因子分析を用いて、構成的妥当性を検証する時、とくに「因子妥当性」と呼ぶ

- （探索的）因子分析が用いられる
  - 項目間の相関関係を分析
  - 背後に潜む潜在因子（ドメインに相当）を探索する
- 解析ソフトのSPSSやSASで計算可能

65

因子分析をひと言で、

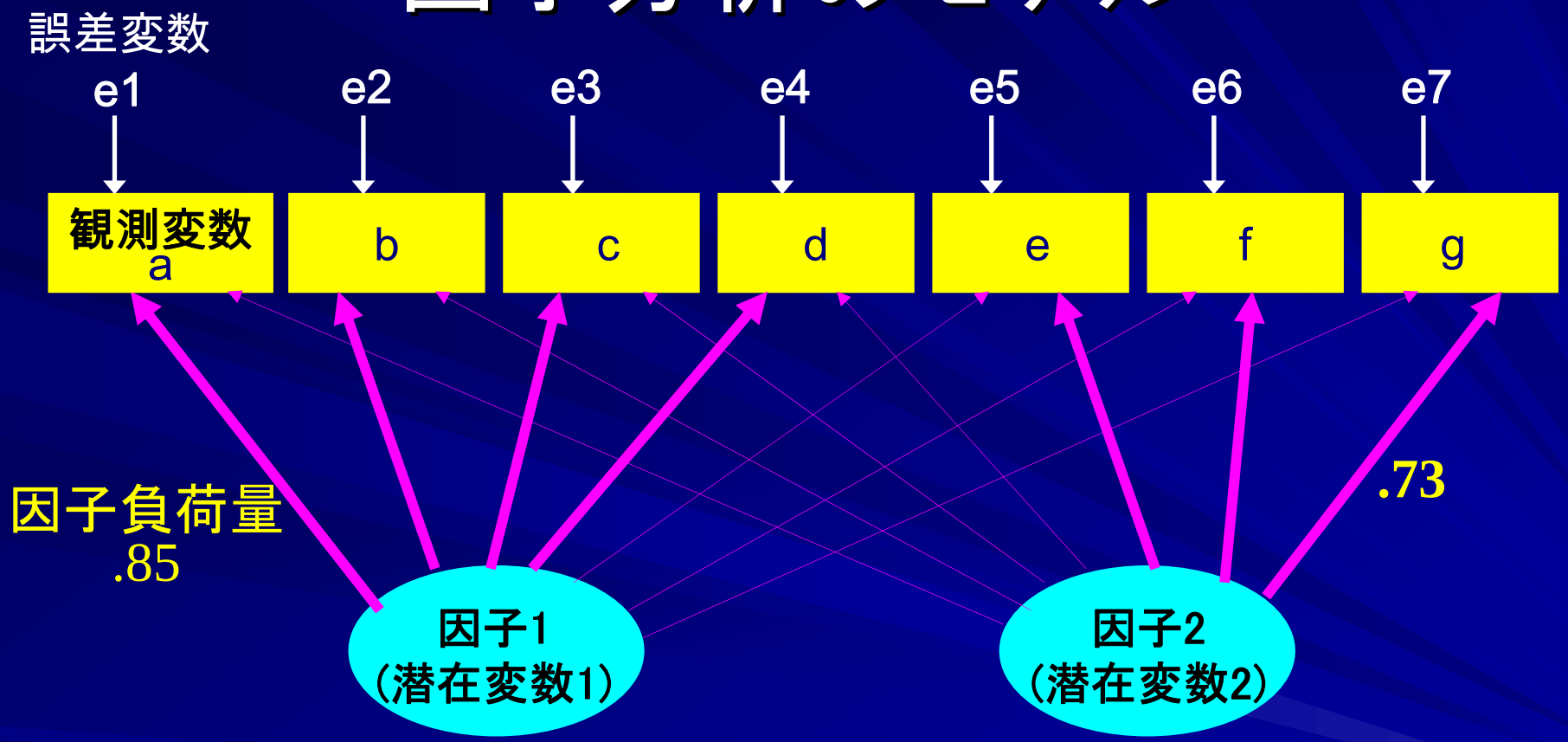
変数（項目）間の相関関係を解析し、

「**因子：潜在変数（latent variable）**」
→　QOL調査票でいうドメインに対応

を見つけ出す分析方法

66

因子分析のモデル



潜在的な変数により変数間に相関が生じていると仮定する

67

因子軸の回転
QOL研究ではバリマックスよりプロマックス

直交回転

- ドメイン相互が無相関（直交）と仮定
- 代表的な直交回転：バリマックス回転

斜交回転

- ドメイン間に相関関係を想定
- QOL研究では適切
- 代表的な斜交回転：プロマックス回転

68

因子分析出力結果（因子負荷量）をどう読むのか？“お作法

行方向に

- 最も高い値にマーク
- それは、0.4以上か?
- すべて因子について0.4未満の項目はないか?

列方向に

- マークした項目群がひと塊になっているか?
- 同じ帰属ドメインを想定した項目群が1つのグループになっているか？
- 別のドメインに入っていないか？

69

因子分析の結果（因子構造）
- 栗原班調査票（進行肺癌）-

表4. 因子構造（プロマックス回転を伴う因子分析）(n=365)

項目	第1因子(活動性)	第2因子(社会性)	第3因子(身体性)	第4因子(精神性)
1. 日常の生活	0.751	0.266	0.476	0.635
2. ひとりで外出	0.817	0.244	0.360	0.560
3. 30分の散歩	0.845	0.246	0.340	0.483
4. 少しの歩き	0.786	0.236	0.444	0.534
5. 階段の昇降	0.792	0.278	0.465	0.498
6. ひとりで風呂	0.670	0.263	0.434	0.388
7. 体の調子	0.587	0.302	0.663	0.651
8. 食欲	0.437	0.232	0.856	0.488
9. 食事	0.411	0.175	0.839	0.469
10. 吐くこと	0.308	0.081	0.396	0.211
11. やせること	0.354	0.124	0.504	0.380
12. 睡眠	0.349	0.322	0.369	0.547
13. 没頭	0.484	0.330	0.380	0.714
14. ストレス	0.527	0.510	0.444	0.768
15. 集中力	0.527	0.458	0.454	0.774
16. 心の支え	0.226	0.163	0.388	0.398
17. 病状に不安	0.231	0.708	0.110	0.413
18. 家族以外の人	0.414	0.513	0.388	0.502
19. 家族に迷惑	0.216	0.729	0.125	0.337
20. 社会生活	0.212	0.822	0.172	0.393
21. 経済的負担	0.203	0.630	0.160	0.308

項目の除外

複数の因子にまたがって関係する項目
どの因子にも関係しない項目

↓

除外候補の項目

↓

除外するかどうかは、他の解析結果と併せて
総合的に判断！

71

構成概念妥当性
-内的整合性-

ドメインに帰属すると想定された複数項目が一つの同じ特性を測定しているか？

クロンバックCronbachの α 係数をドメインごとに計算

ドメイン内の項目が互いに相関していることを検討するための指標

伝統的に、0.6もしくは0.7以上の値が望ましい

0.9を超えて1に近い場合

すべての項目が同じものを測定：無駄に多くの項目を測定の可能性

SPSSやSASで計算可能

72

構成概念妥当性
-内的整合性-

栗原班調査票(進行肺癌)

ドメイン	項目数	クロンバックのα係数
活動性	5	0.90
身体性	5	0.80
精神性	4	0.81
社会性	5	0.82

FACT-G／Sp(前立腺癌、子宮癌、他)

Sub Scale	Total Items	Cronbach's alpha
Physical Well-Being	7	0.82
Social/Family Well-Being	6	0.81
Emotional Well-Being	6	0.83
Functional Well-Being	7	0.86
Spiritual Well-Being (items 1-8; meaning/peace)	8	0.89
Spiritual Well-Being (items 9-12; faith)	4	0.82

73

基準関連妥当性
-臨床的妥当性-

臨床的データとの関係を調べる
回帰分析
ドメインスコアを目的変数、臨床パラメータを説明変数
栗原班（活動性、身体性）：進行肺癌

説明変数[a]	活動性		身体性	
	回帰係数	p値	回帰係数	p値
PS	-0.38	<0.001	-0.16	0.032
臨床病期	-0.14	0.179	-0.06	0.418
性別	0.31	0.010	0.13	0.143
年齢	-0.08	0.422	0.02	0.779
体重減少	-0.40	<0.001	-0.42	<0.001
アルブミン	-0.09	0.485	-0.25	0.005
LDH	-0.25	0.131	-0.23	0.047

[a] 説明変数: PS(2/1/0), 臨床病期(IV/IIIB), 性別(男/女), 年齢(≧65/≦64歳), 体重減少(≧5%/<5%), アルブミン(<3.5/≧3.5 g/dl), LDH(≧500/<500 IU/l)

74

お作法 その4

計量心理学的解析

- 実施可能性/項目分析　feasibility/item analysis
- 信頼性　reliability
- 妥当性　validity
- 感度/反応性　sensitivity/responsibility

75

感度/反応性

対象者の状態変化→QOLスコアも変化？

- 治療などの外的要因（治療効果）
- 疾病の悪化

ドメインスコアの変化量を目的変数

分散分析、共分散分析（前値を共変量）

76

感度/反応性

進行肺癌化学療法1コース（栗原班調査票）

分散分析

要因：Performance status (PS)の変化

Table 7. Sensitivity of the domain scores to changes in PS during the first course of treatment

		Change of mean score			
	n	Functional	Physical	Mental	Psychosocial
Deteriorated PS	49	-0.76	-0.60	-0.48	0.00
Unchanged PS	248	-0.37	-0.26	-0.12	0.09
Improved PS	25	-0.23	-0.06	-0.10	0.07
ANOVA[a]		*P* = 0.015	*P* = 0.005	*P* = 0.008	*P* = 0.661

[a] One-way analysis of variance based on F statistics: F (2, 319)

お作法のまとめ

実施可能性	項目欠測、記入時間、頻度分布		欠測割合、ヒストグラム
信頼性	test-retest再現性		級内相関係数、κ 係数
妥当性	内容的		--------
	構成概念	収束的	Multitrait scaling analysis
		弁別的	
		因子	因子分析
		内的整合性	Cronbach's α
	基準関連	予測	Cox回帰、Logisitic回帰
		臨床的	Known group comparison、回帰分析
感度	患者の状態変化との関係		分散分析、共分散分析

78

臨床試験のエンドポイントとしてのQOL評価

臨床試験のエンドポイントとして
-がん領域では-

QOL調査票データ

→　副次的評価項目として

Quality Adjusted Life Years (QALYs)

Clinical Benefit Response (CBR)　膵がん

→　主要評価項目として

生存の質と量の総合評価:
QALYs（Quality Adjusted Life Years）について

質で調整した生存年

単純な生存の比較ではない

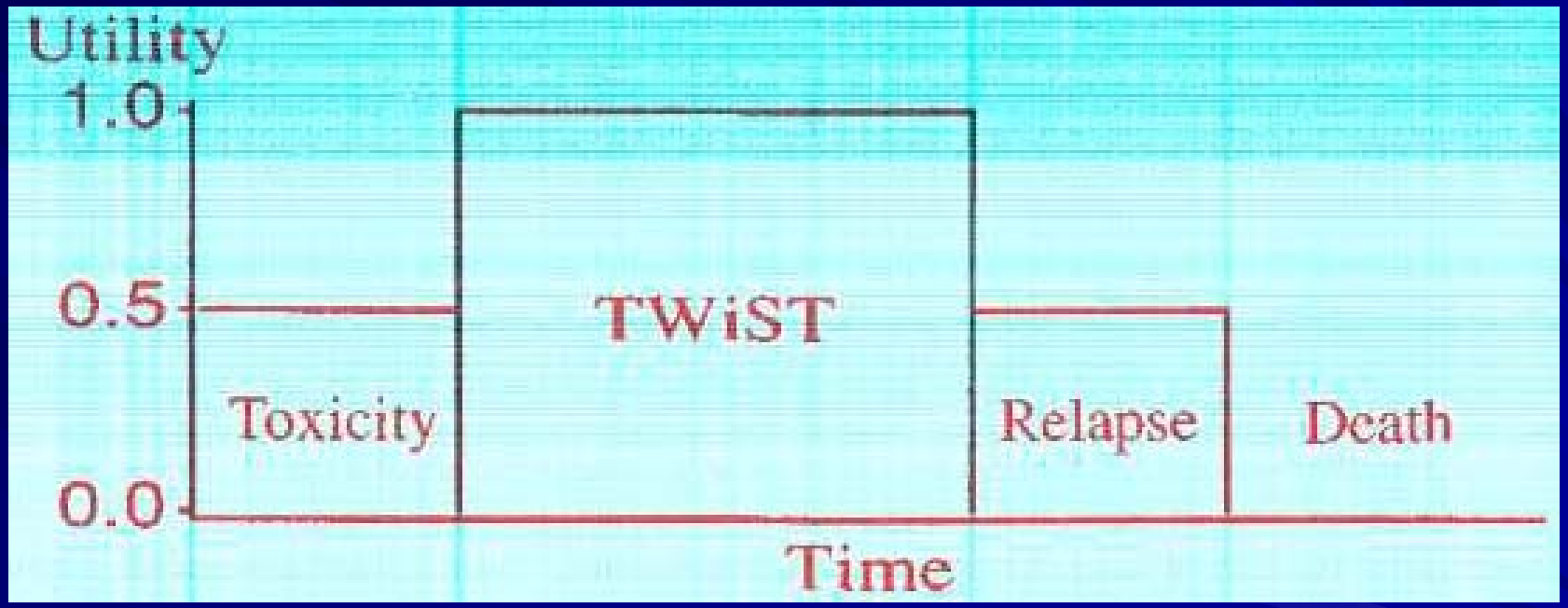
生存の量と質を統合した指標

術後乳癌患者に対する補助療法や内分泌療法の試験が最初

81

Q-TWiST

Quality-adjusted Time Without Symptoms of disease or Toxicity of treatment



Q-TWiST

$= (U_{TOX} \times TOX) + (U_{TWiST} \times TWiST) + (U_{REL} \times REL)$

Uはその状態での生存時間に掛ける重み

82

Clinical Benefit Response

膵癌に対するGemcitabine
米国FDAによる認可の際用いられた

プライマリーエンドポイントとして
Pain（Analgesic）とPSの合成変数
（J.Clin Oncol 1994）

83

QOL調査計画
-キーポイントは3つ-

1) 調査票（あるいは指標）の選択

2) 調査期間、調査時点およびポイント数

3) 評価ウィンドウの設定

84

調査票の選択

研究目的、対象疾患、実施可能性（項目数）を考慮して決定

- 包括的調査票＋疾患特異的　が一般的
- 疾患特異的な症状に焦点を絞る
  →患者負担を考慮、実践的かつ現実的な場合あり

例：進行・再発胃癌、化学療法

- EORTC QLQ-C30+STO22（日本語版）
- FACT-GC（日本語版）
- 調査項目の追加も有用
  例：ダンピング症候群（食後のめまい、発汗など）

85

調査の時点／期間

対象疾患、治療法、研究目的によって異なる

ベースライン（治療開始前）データに加えて、

- 例1：癌化学療法の制吐剤
  - 症状が発現する数日から1週間
- 例2：進行癌に対する癌化学療法
  - 化学療法中と終了後
  - 治療中→各コース開始前
- 例3：術後乳癌患者対象の補助療法
  - 通常、研究期間が1年を超えるため
  - 3～6ヶ月毎くらいの調査が現実的

86

評価ウィンドウの設定1

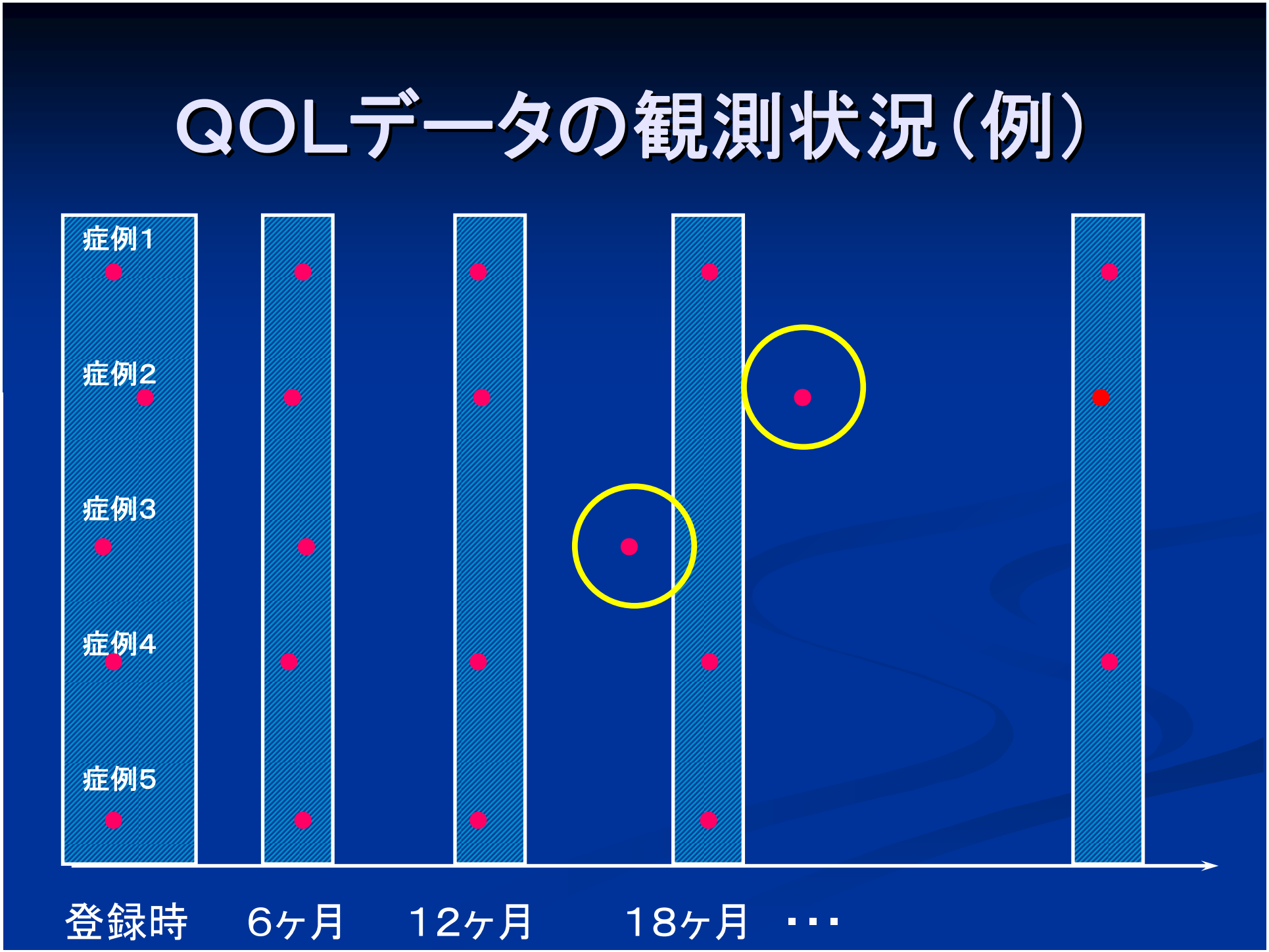
例えば、術後乳癌患者対象に補助療法AとBを比較する臨床試験を計画

- 研究期間は5年
- 治療開始時 + 6ヶ月毎の調査

実際の試験では、設定したスケジュールからずれて観測されるデータが必ず出てくる

- 治療開始後17ヶ月目のデータは?　採用?
- 20ヶ月目のデータは?　非採用?

87


QOLデータの観測状況（例）
症例1
症例2
症例3
症例4
症例5
登録時
6ヶ月
12ヶ月
18ヶ月 ・・・

評価ウィンドウの設定2

調査予定日の前後何日（週）までを解析に用いるのか?

データの採用幅:“評価ウィンドウ”を研究計画時に決めておくことが重要

- 比較的安定した疾患を対象、3ヶ月に1回の調査
  2～3週程度
- 進行癌の化学療法、各コース直前の測定
  3日程度（長くとも1週間）

89

QOLデータの解析は参考資料に

Take home message 1

- QOLとは何か？その定義を理解
- QOL測定には患者自記式調査票が必要
  - 開発には、翻訳と新規作成
  - 標準的な手順あり
- 調査票のバリデーション
  - 実施可能性、信頼性、妥当性、感度
  - 計量心理学的な解析
- 臨床試験のエンドポイント：CBR、QALYs

91

Take home message 2

QOLデータの回収

記入時期ズレ、データ欠測をなくす努力

研究デザイン・実施体制の工夫

調査時期前に担当医、CRCにリマインド

QOLデータの解析

欠測値の問題を除けば他のデータ解析と同じ

データ欠測への対応が重要な鍵

参考文献1

教科書

- 池上,他(編集). 臨床のためのQOL評価ハンドブック, 医学書院, 2001
- 漆原(監修). 新QOL調査と評価の手引き. メディカルレビュー社, 2001
- 福原,数間(監訳). QOL評価学-測定, 解析, 解釈のすべて-(P Fayers, D Machin著). 中山書店, 2005

その他

- 下妻晃二郎, 森田智視. QOL評価. 癌治療の新たな試み 新編III. 医薬ジャーナル社, 714-736. 2005.
- 大橋靖雄. 臨床研究で用いられる健康関連QOL尺度に要求される計量心理学的特性. 臨床透析 1997; 13: 1083-1090.

Textbooks

- Spilker B, ed, Quality of Life and Pharmacoeconomics in Clinical Trials, second edition. Lippincott-Raven, Philadelphia, 1996
- Staquet MJ, Hays RD, Fayers PM. Quality of Life Assessment in Clinical Trials. Oxford, Oxford University Press, 1998
- Mesbah M, Cole BF, Lee MLT eds. Statistical Methods for Quality of Life Studies. Boston: Kluwer Academic Publications, 2002.
- Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. New York: Chapman & Hall/CRC, 2002.

93

参考文献2

Articles

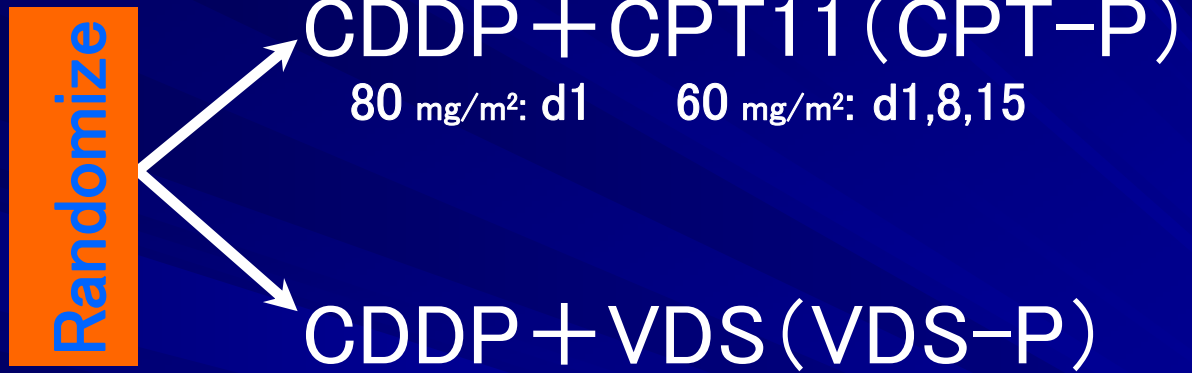
- Machin D. Assessment of quality of life in clinical trials of the British Medical Research Council. J Natl Cancer Inst Monogr 1996; 20: 97-102.
- Moinpour CM, Feigl P, Metch B, et al. Quality of life end points in cancer clinical trials: Reviews and recommendations. J Natl Cancer Inst 1989; 81: 485-495.
- Nayfield SG, Ganz PA, Moinpour CM, et al. Report from a National Cancer Institute (USA) Workshop on quality of life assessment in cancer clinical trials. Qual Life Res 1992; 1: 203-210.
- Cella DF, Tulsky DS. Quality of life in cancer: definition, purpose, and method of measurement. Cancer Invest 1993; 11: 327-336.
- Osoba D. Lessons learned from measuring health-related quality of life in oncology. J Clin Oncol 1994; 12: 608-616.
- Landis JR Koch GG.The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174.
- Johnson JR, Temple R. Food and Drug Administration requirements for approval of new anticancer drugs. Cancer Treat Rep. 1985; 69: 1155-1159
- HA Burris, 3rd, MJ Moore, J Andersen, et al. Improvements in survival and clinical benefit with gemcitabine as first- line therapy for patients with advanced pancreas cancer: a randomized trial. J Clin Oncol 1997: 15: 2403-2413

計量心理学的解析

- Coste J, Fermanian J, Venot A. Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. Stat Med. 1995; 14: 331-345.
- 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 Inst 1993; 85: 365-376.
- Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy Scale: Development and validation of the general measure. J Clin Oncol 1993; 11: 570-579. Aaronson NK, Bullinger M, Ahmedzai S. A modular approach to quality-of-life assessment in cancer clinical trials. Recent Results Cancer Res 1988; 111: 231-249.
- Kaasa S, Bjordal K, Aaronson NK, et al. The EORTC Core Quality of Life Questionnaire (QLQ-C30): Validity and reliability when analyzed with patients treated with palliative radiotherapy. Eur J Cancer 1995; 31A: 2260-2263.

参考資料

事例：抗癌剤第Ⅲ相試験



前治療のない非小細胞肺癌患者

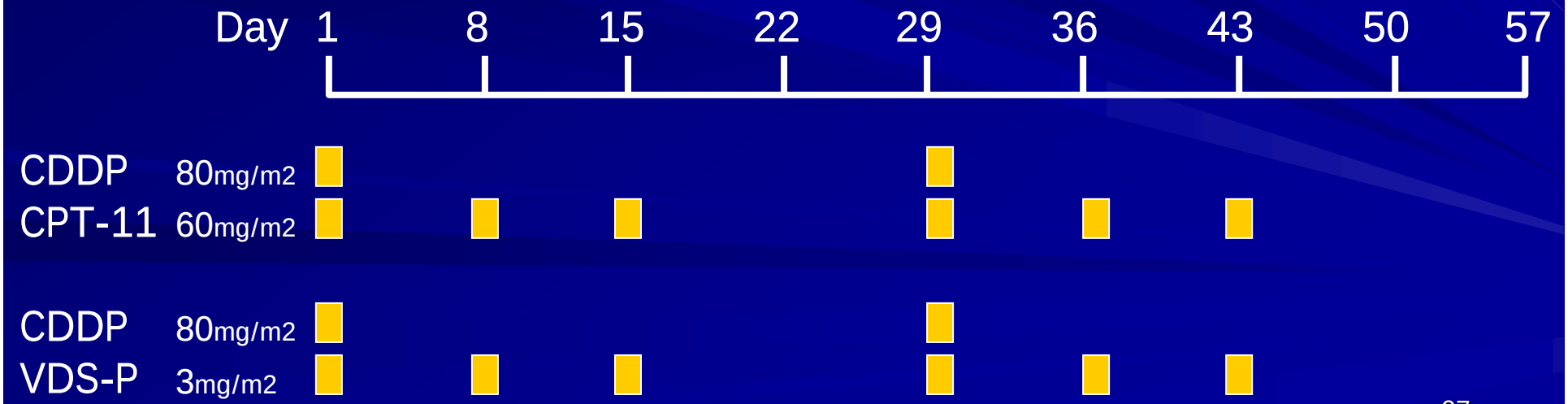
Stage：ⅢB／Ⅳ、

ECOG PS：0～2

QOL:　副次評価項目の一つ

試験治療

- 1コース4週
- 各コースdays 1,8,15に抗癌剤投与、1週休薬
- 少なくとも2コース以上施行
- Progressive diseaseになるまで



97

QOL調査票(QOL-ACD)

記入年月日	199　年　月　日

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。
この数日間のあなたの状態にあてはまると思われる番号に○をつけて下さい。
（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。

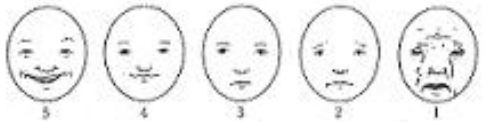
記入例：1　2　3　④　5

【この数日の間】

1. 日常の生活（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. 少し歩いてもつらいと思いましたか。　1 2 3 4 5　非常につらかった ― 全く問題なかった
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
6. ひとりで風呂にはい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
7. 体の調子はいかがでしたか。　1 2 3 4 5　非常に悪かった ― 非常に良かった
8. 食欲はありましたか（食べたいと思いましたか）。　1 2 3 4 5　全くなかった ― 非常にあった
9. 食事がおいしいと思いましたか（一口でもおいしく感じましたか）。　1 2 3 4 5　非常にまずかった ― 非常においしかった
10. 吐くことがありましたか。　1 2 3 4 5　よく吐いた ― 全く吐かなかった
11. やせましたか。　1 2 3 4 5　非常にやせた ― 全くやせなかった
12. よく眠れましたか。　1 2 3 4 5　全く眠れなかった ― よく眠れた
13. 何かに没頭（熱中）することができましたか。　1 2 3 4 5　全くできなかった ― よくできた

（裏ページへ）

14. 日々のストレス（いらいら）はうまく解消できましたか。　1 2 3 4 5　全くできなかった ― うまくできた
15. 集中力が落ちたと感じましたか。　1 2 3 4 5　強く感じた ― 全く感じなかった
16. 何か心の支えになるものによって勇気づけられていますか。（家族，知人，宗教，趣味など）　1 2 3 4 5　全くない ― 強く勇気づけられている
17. あなたの病状に不安を感じましたか。　1 2 3 4 5　強く感じた ― 全く感じなかった
18. 家族以外の人と接するのが苦痛でしたか。　1 2 3 4 5　非常に苦痛だった ― 全く問題なかった
19. あなたが治療をうけていることで家族に迷惑をかけていると思いますか。　1 2 3 4 5　強く思っている ― 全く思わない
20. あなたの将来の社会生活について不安を感じますか。　1 2 3 4 5　強く感じる ― 全く感じない
21. 病気による経済的な負担が気になりますか。　1 2 3 4 5　非常に気になる ― 全く気にならない
22. ここ数日間の状態に相当する顔の番号に○をつけて下さい。



☆最後にもう一度，つけ落としがないか確認して下さい。

（担当医師・看護婦確認欄）

1. 患者さんに記入いただいた内容に記入漏れがないか確認すること。
2. 全部記入されていれば，患者さんより本調査票を受け取り，下欄にサインをした上で，次回調査まで保管しておくこと。

確認年月日	199　年　月　日
確 認 者	
患者氏名	

QOL調査票

- 栗原班QOL－ACD
  - 全22項目の自記式調査票
    - 活動性ドメイン　（項目1～6）
    - 身体性ドメイン　（項目7～11）
    - 精神性ドメイン　（項目12～16）
    - 社会性ドメイン　（項目17～21）
    - グローバルQOL　（項目22：フェーススケール）

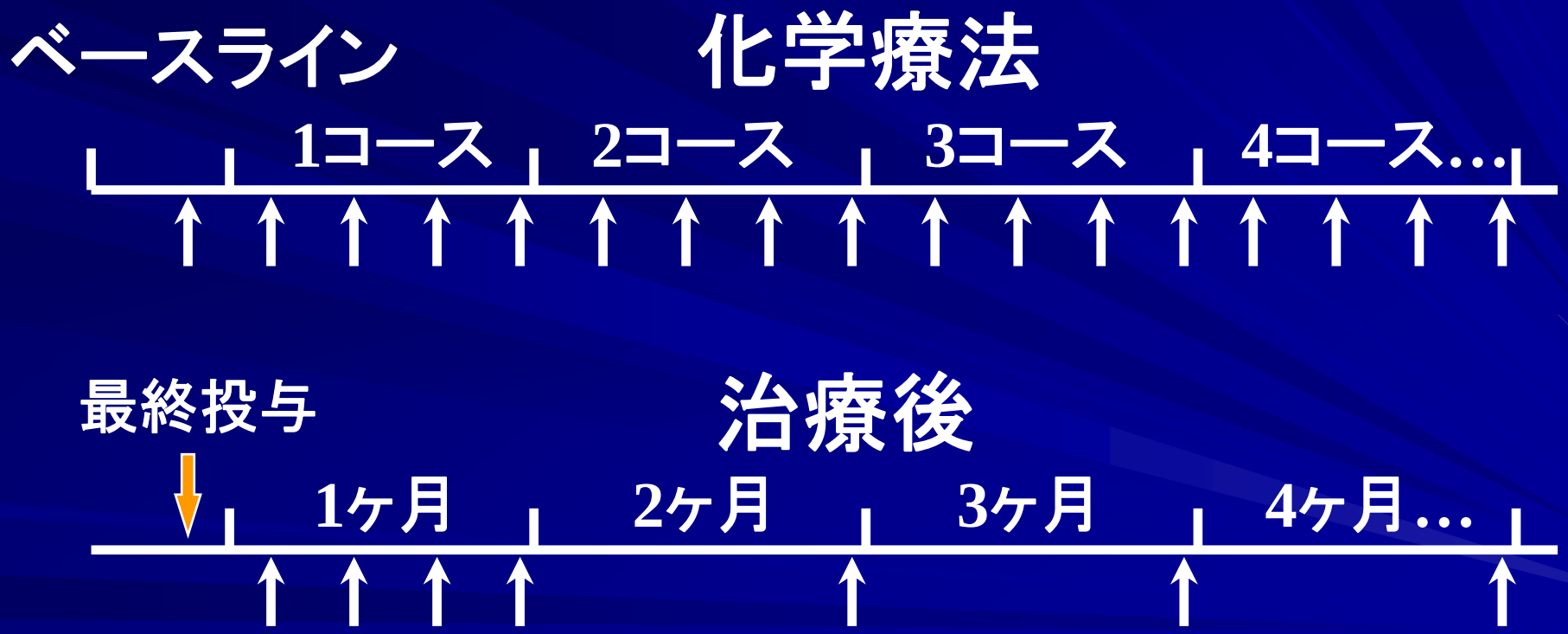
解析目的とデータ

抗癌剤副作用による短期的な影響とともに
抗腫瘍効果／病態悪化のQOLに与える影響を
評価したい

– 治療前→1コース後→2コース後 の経時的変動

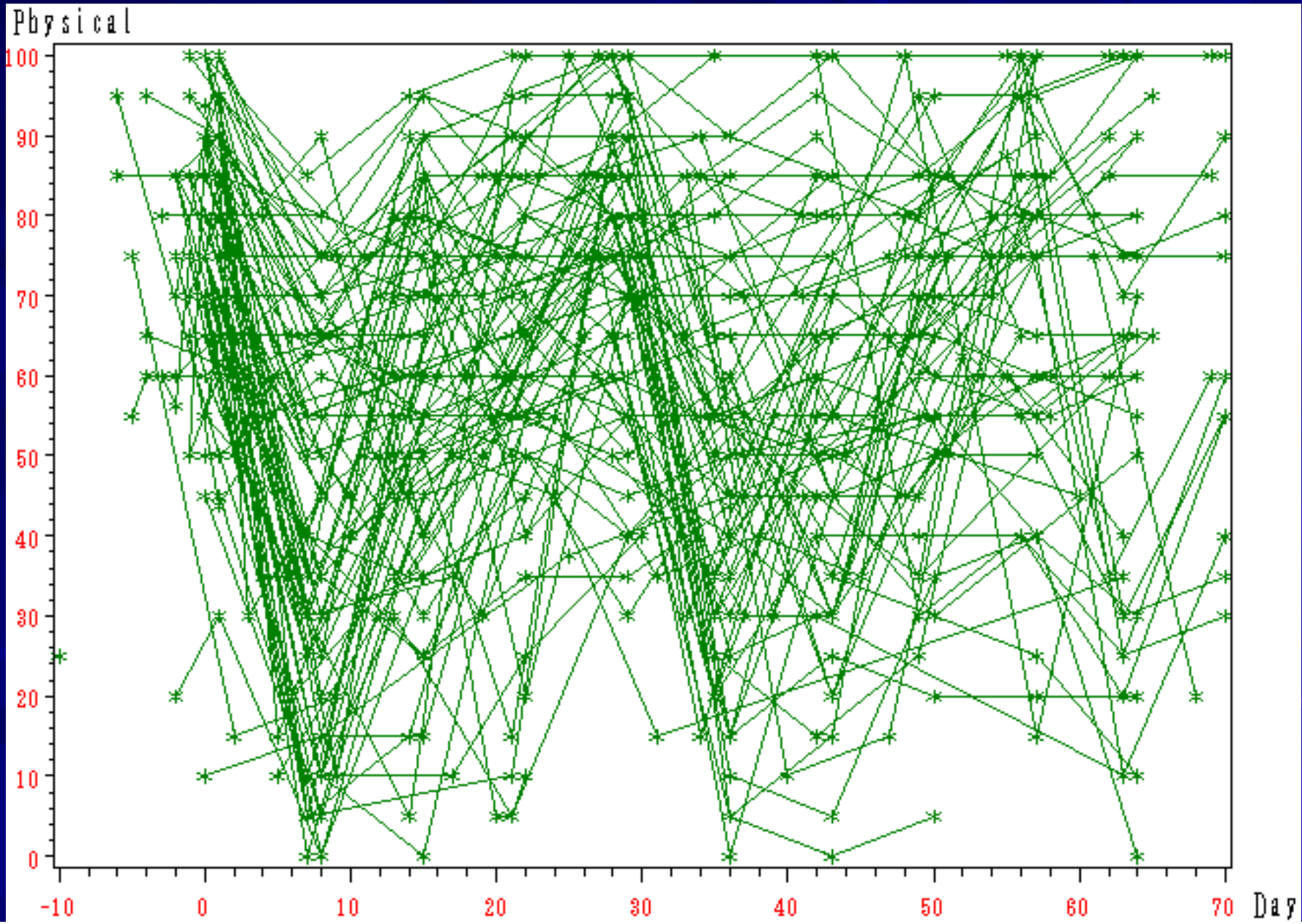
100

QOLデータ収集計画

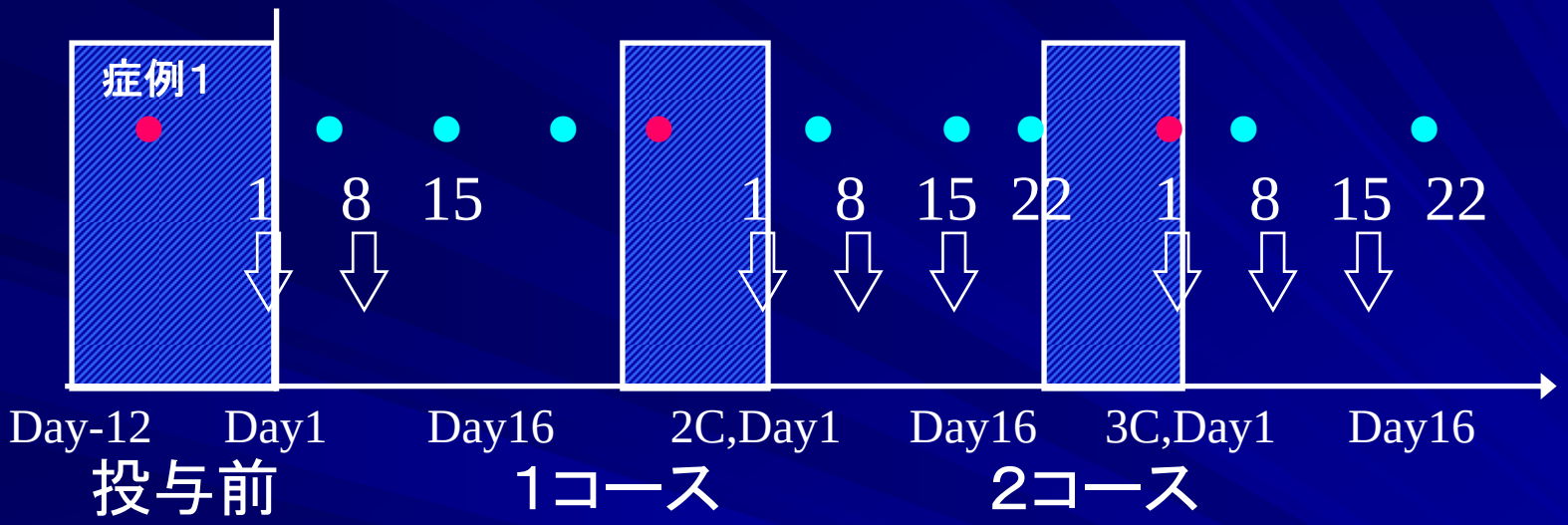


101

身体性ドメインスコアの推移



ウィンドウの設定



治療前→1コース後→2コース後 の経時的変動

次コース開始前7日間の評価ウィンドウ

この試験では毎週データ収集だが→患者負担も考慮すると各コース前が最適

103

肺癌研究で議論したポイント

- 抗癌剤投与当日の投与直後のデータ
  - 点滴による疲れなどによる一時的なQOL低下
  - 本来の目的からずれているため、投与後に記入されたデータは解析から除外

Note:　記入時点が投与前後のどちらか？

→　調査票に記入欄あるいはCRCの調査項目に

 無断転載禁止　http://www.icrweb.jp/icr/

肺癌研究で議論したポイント

- 後治療開始後に収集されたデータ
  - プロトコル治療終了後に施行される後治療（放射線、他の化学療法）の影響を否定できない
  - 生存時間などの解析では、RCTの場合、ITTの観点から許容、
  - しかし、QOLなどの経時的データでは解釈の困難さ

→後治療開始後のデータは、解析から除外

105

CRCの役割

- 従来の臨床試験では、
  - 収集データ：遵守状況を見てウィンドウを決定することが行われていた
- 現在では、
  - CRC（clinical research coordinator）が研究の質管理に関与
  - プロトコル策定時にウィンドウ設定
  - 正確な測定時点、高いデータ回収率

106

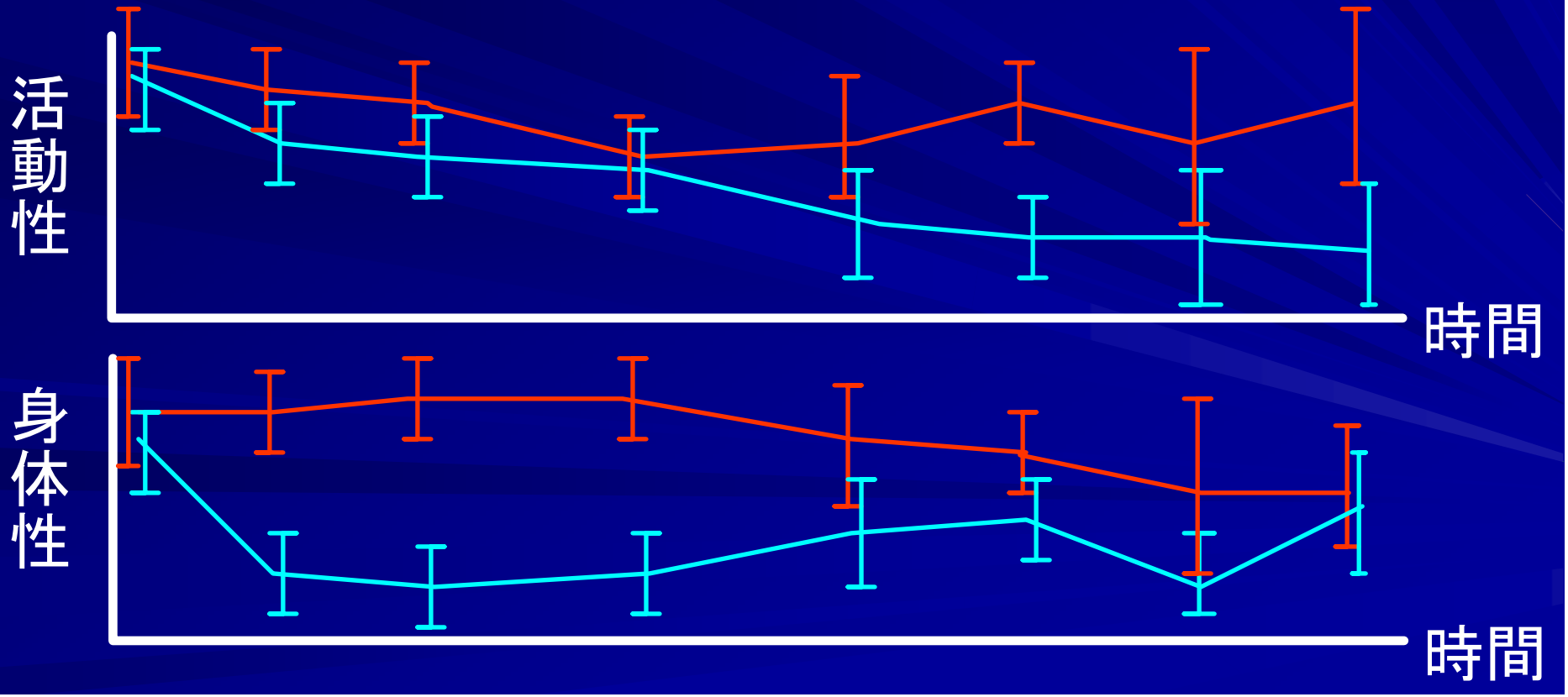
QOLデータ解析、なぜ厄介?

- 多時点・多項目で測定
  - 検定の多重性
  - 第一種の過誤の増大
- 欠測値
  - データ欠測が患者のQOLと関連
  - 状態の良い患者ほど長期に追跡
  - 解析結果にバイアス
- 結果の提示方法
  - 臨床的意味と解釈の難しさ

107

QOLスコアの変動（例）

- 少なくとも3時点（治療前／中／後）
- 複数のドメイン（活動性、身体性…）



データ解析方法
大きく3つのアプローチ

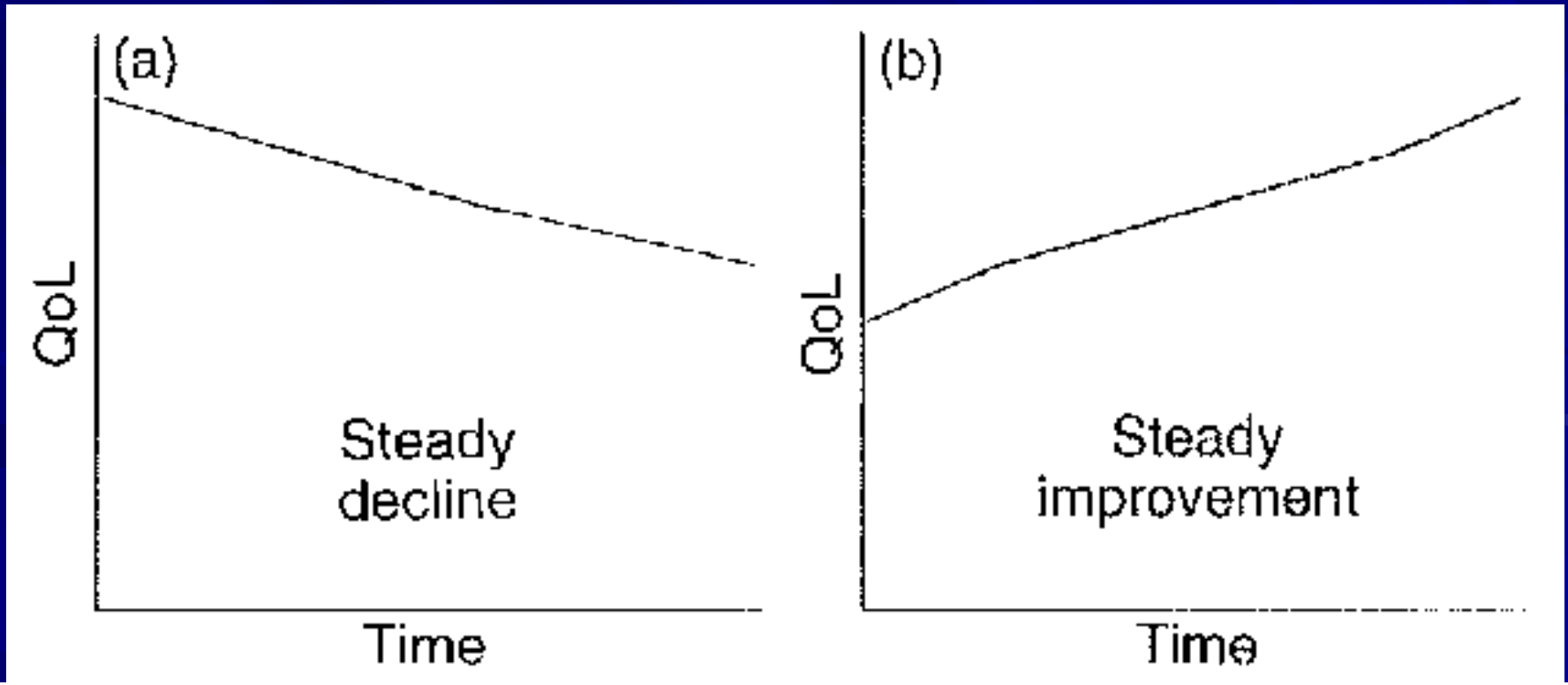
1)　意味ある測定時点をあらかじめ設定
　　例：治療開始前、治療中、治療終了後4ヵ月

2)　経時的な情報を一つの要約指標に集約
　　例：傾き、AUC

3)　測定されたデータすべてを解析

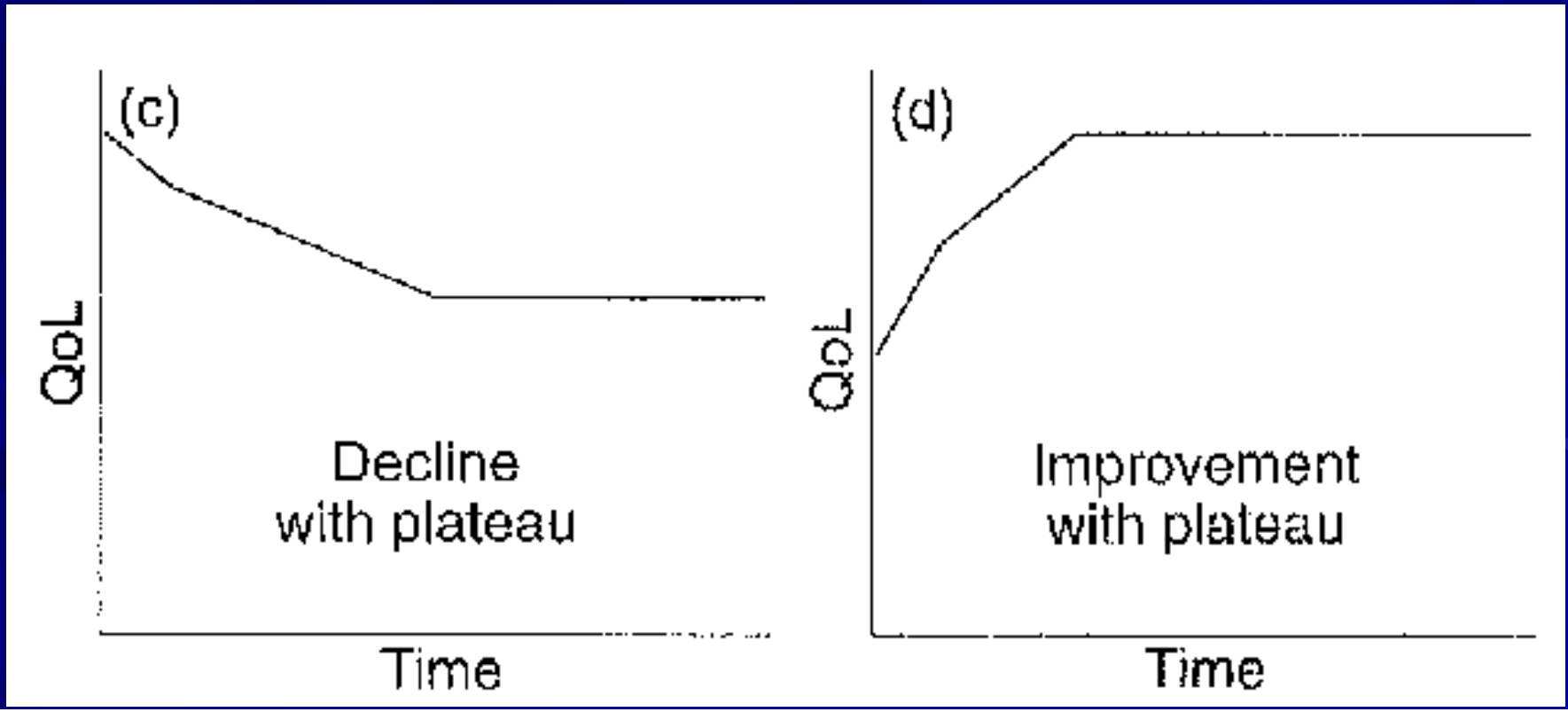
要約指標1

単調な変動
・傾き、・ベースラインと最終測定値の差



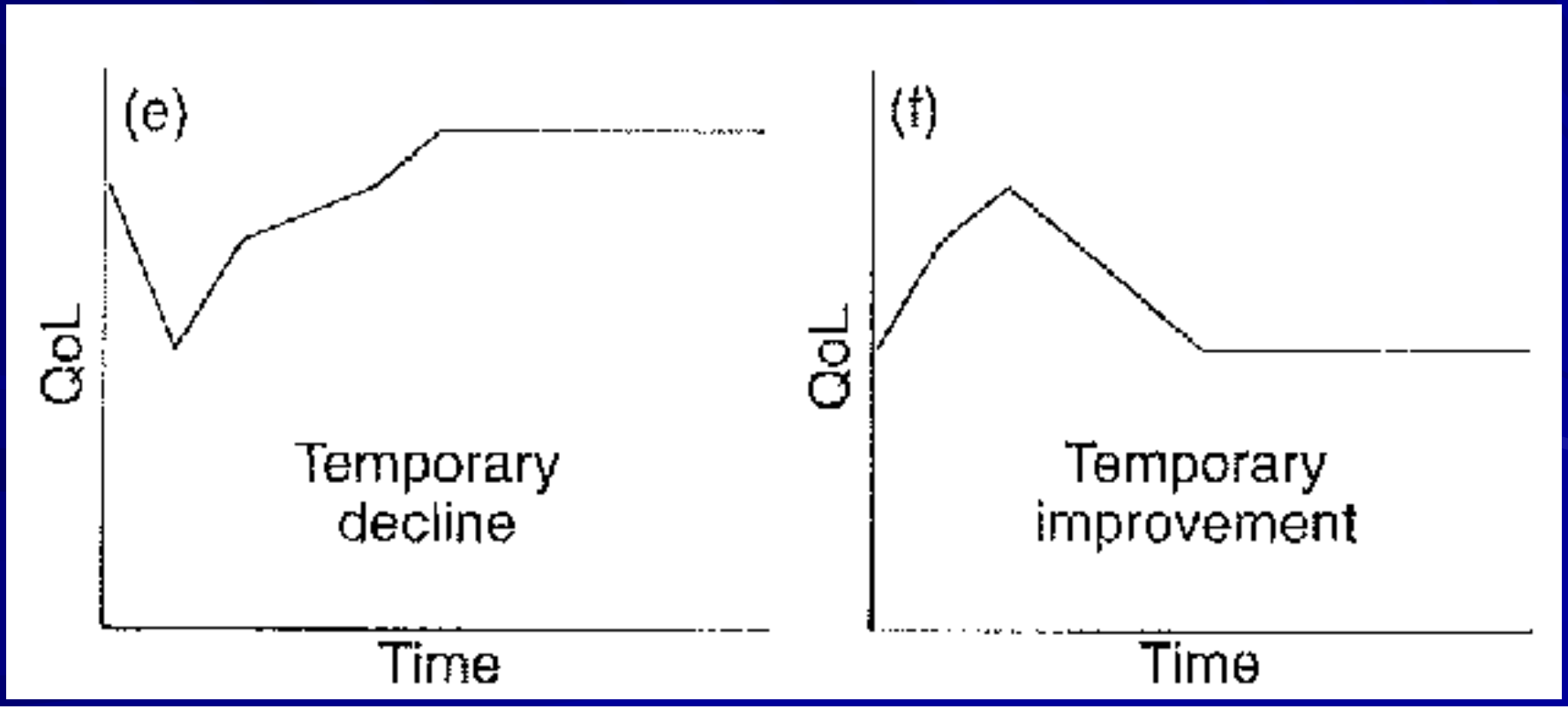
要約指標2

飽和状態
- ピークに達するまでの時間
- 治療後の測定値の平均



要約指標3

一時的な変動
- AUC



解析方法 3）

経時測定データに対する解析方法

- 経時測定データに対する分散分析
- 混合効果モデル(Mixed Effects Model)
- GEE（generalized estimation equation）

どの解析方法?

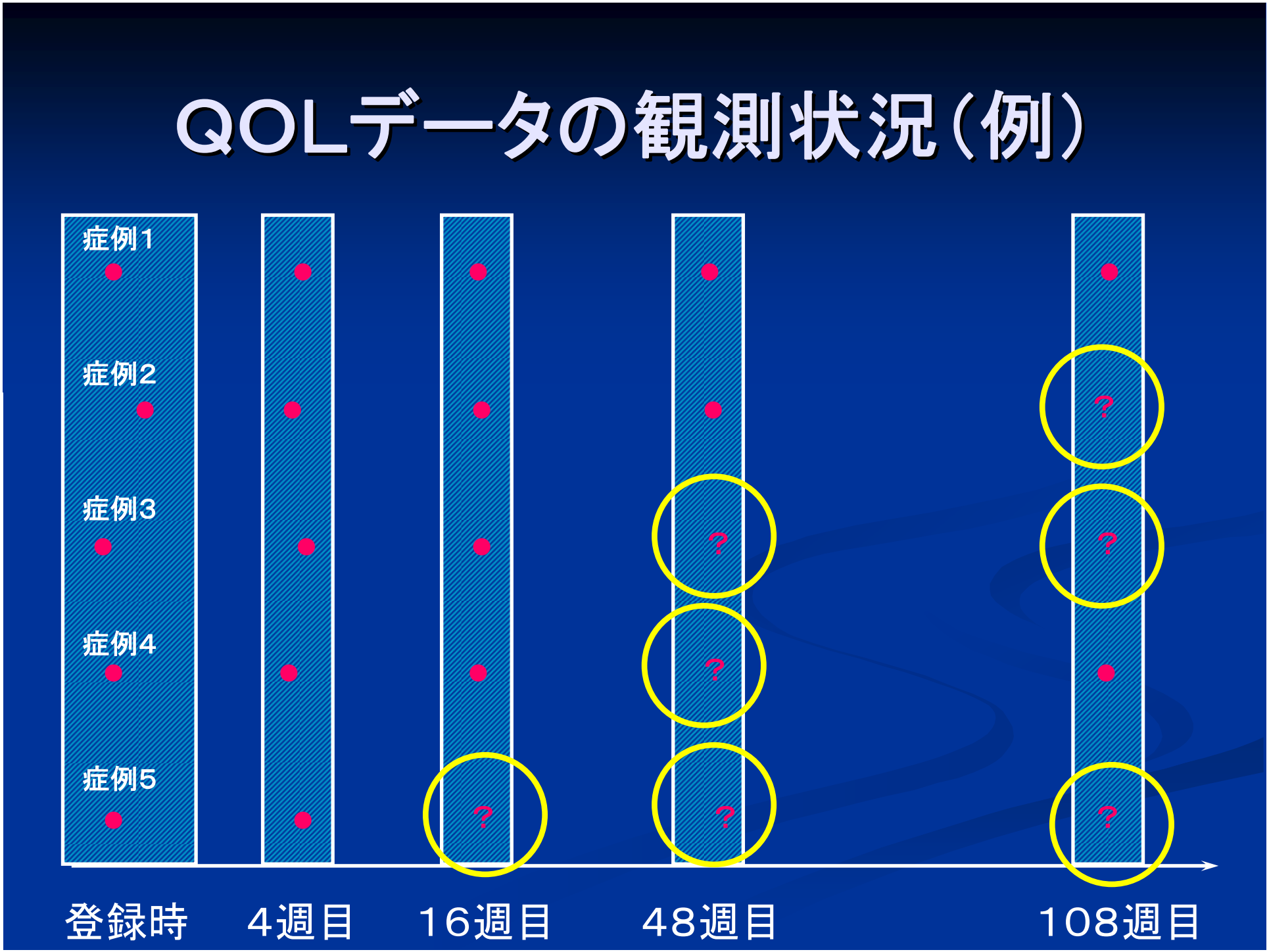
解析のシンプルさ･･･解釈のしやすさ
- 時点の絞込み
- 要約指標

努力して測定した経時データ･･･経時的なプロファイルの評価、感度／検出力↑
- 混合効果モデル、GEE
- データ解析ソフトウェアの充実

何れの方法も欠測データが少ない場合はOK

しかし、現実には･･･欠測への対処策が必要な場合がある

114


QOLデータの観測状況（例）
症例1
症例2
症例3
症例4
症例5
?
登録時
4週目
16週目
48週目
108週目

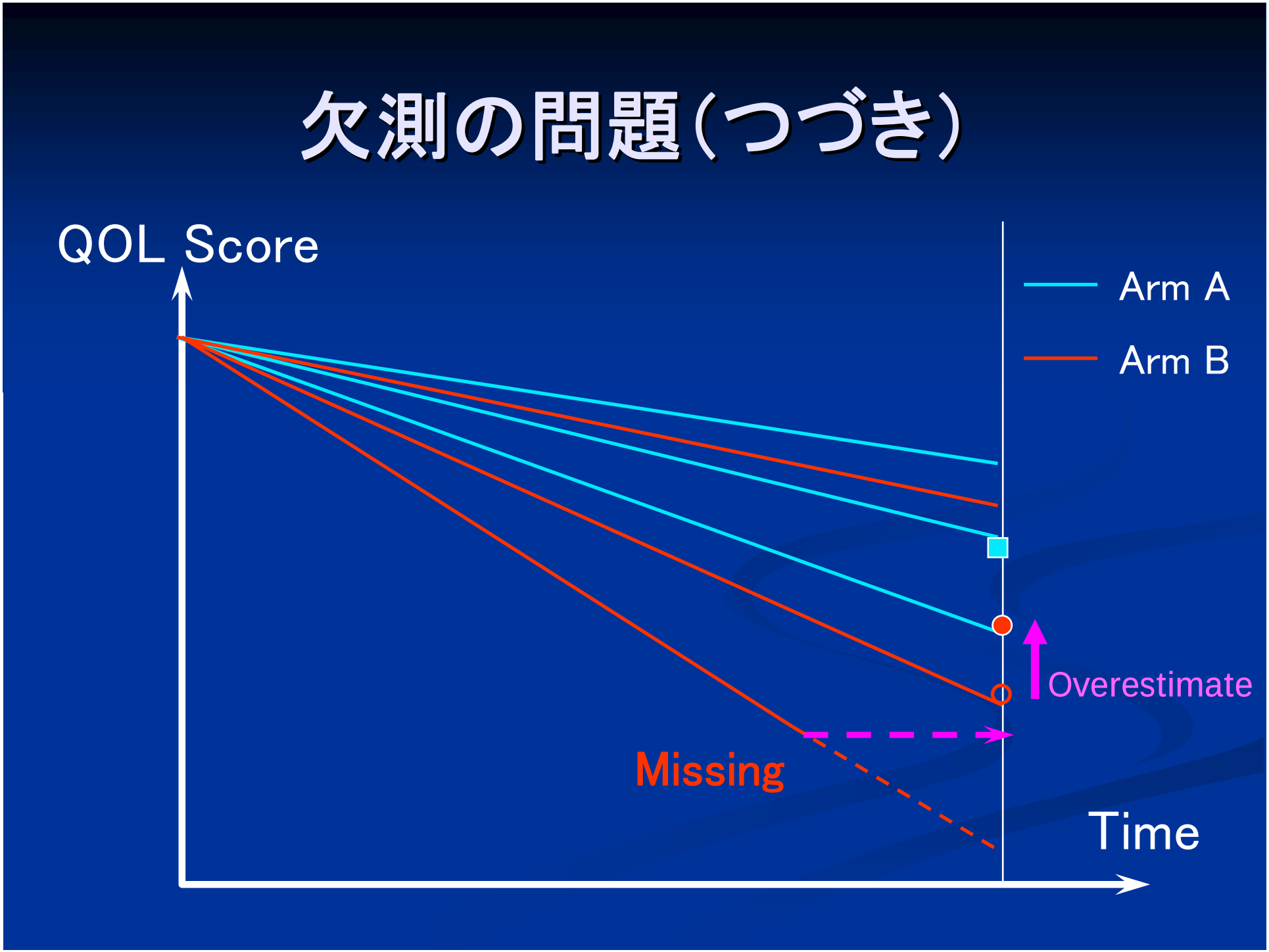
欠測値が存在すると一般的なデータ解析ソフトウェアでは

- 時点ごとの平均値の算出
  - 欠測値を除いて計算
- 2時点間の相関
  - 特別な設定なしではペアワイズで削除
- 経時測定データの解析
  - リストワイズで（1時点でも欠測があればその症例データまるごと）削除

欠測の問題

- 検出力（or 精度）の低下
  - データ数の減少
  - 精度の減少（信頼区間が広くなる/有意になり難い）
  - 対応は可能：集積症例数を多めに
- 選択バイアス（解析結果に偏り）
  - 状態の良好な症例ほど長期フォロー
  - 観察されたデータのみに基づく解析にはバイアス
  - 予後の悪い疾患を対象・・・シリアスなバイアス

欠測の問題（つづき）



生じてしまったら実際問題どうする？

観測された患者集団とデータが欠測した患者集団との比較

例：QOL質問紙調査における無回答者

- 「回答割合は@@%でした」だけでなく
- 回答者と無回答者で、性・年齢などの背景因子の比較
- 少なくとも背景因子の分布の比較を考察しておく

具体的な解析方法については統計家に相談