Patient-Reported Outcomes: End Points for Ovarian Cancer

Lari Wenzel, Ph.D.

Karen Basen-Engquist, Ph.D.

Presentation Objectives

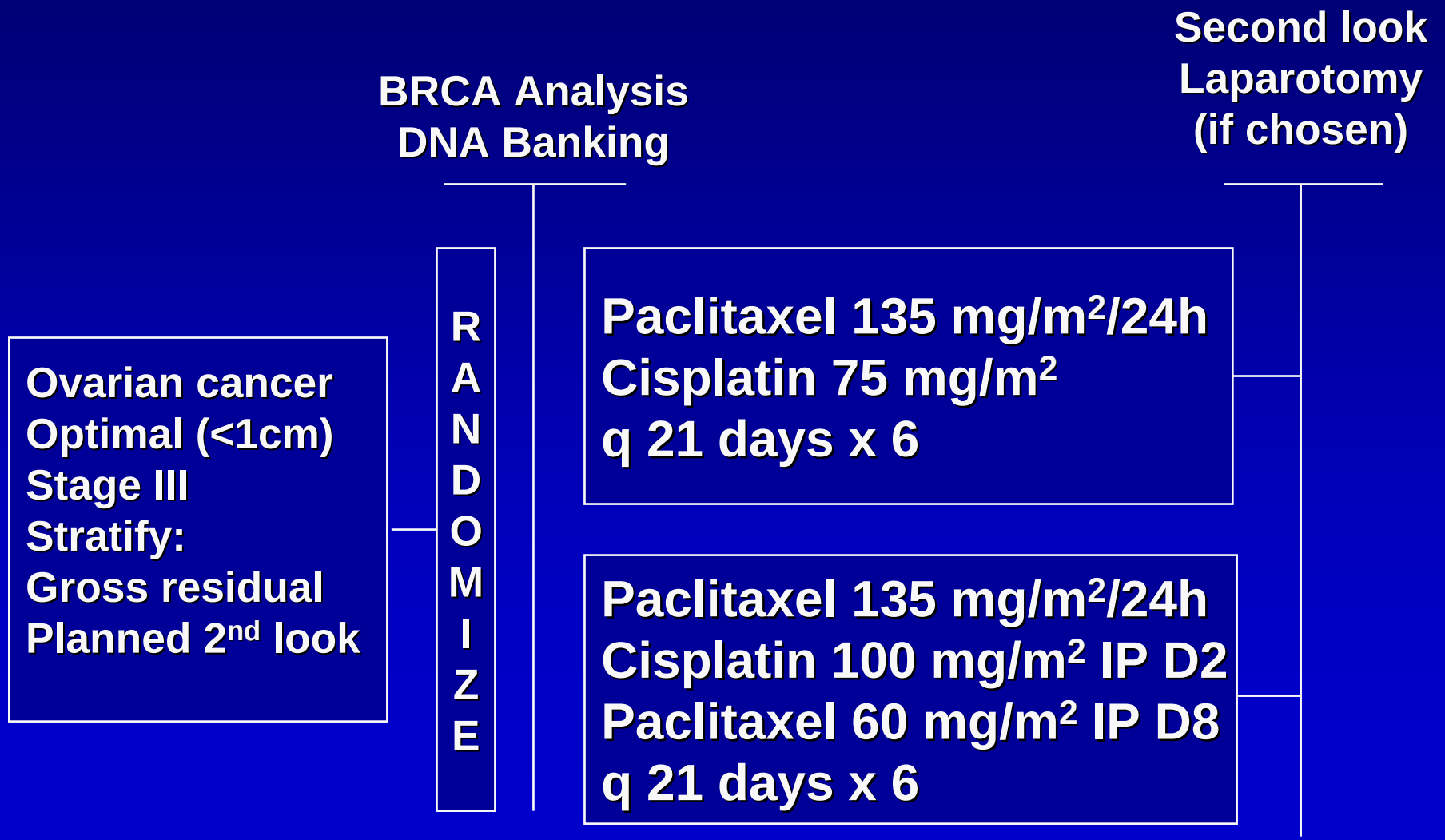
- Illustrate significant PRO end points in ovarian cancer trials
- Identify PRO measures ready for incorporation
- Highlight PRO issues for further study

Why are Patient-Reported Outcomes Important in Ovarian Cancer?

- Ovarian cancer treatments should be evaluated for ability to improve patient functioning and reduce symptoms
- Treatment-related side effects must also be assessed
- Symptoms and function are best measured by asking patients directly
- Examples from recently completed Phase III trials

GOG #172

Armstrong et.al. Abs #803, ASCO 2002



GOG #172: Survival

	Regimen 1 Intravenous	Regimen 2 Intraperitoneal
Progression-free	**18.3 mos**	**23.8 mos**
Overall Survival	**49.5 mos**	**66.9 mos**

Results of GOG 172

Armstrong et al., 2006

- The IP regimen used higher and more frequent dosing than the IV regimen
- Toxicities were greater on the IP arm
- Fewer patients on the IP arm were able to complete 6 cycles of therapy
- A statistically significant improvement in PFS and OS for patients in the IP arm
- **The 65.6 month median survival on IP is the longest survival reported to date from an advanced OC randomized trial**

Consensus: 2005

- The toxicities, inconvenience and cost of IP therapy are justified by the improved survival seen with this treatment
- New, targeted therapies are likely to be more effective in patients who have an excellent response to chemotherapy
- While we work to improve the tolerability and toxicities of IP therapy, it remains the most effective means of treating ovarian cancer today

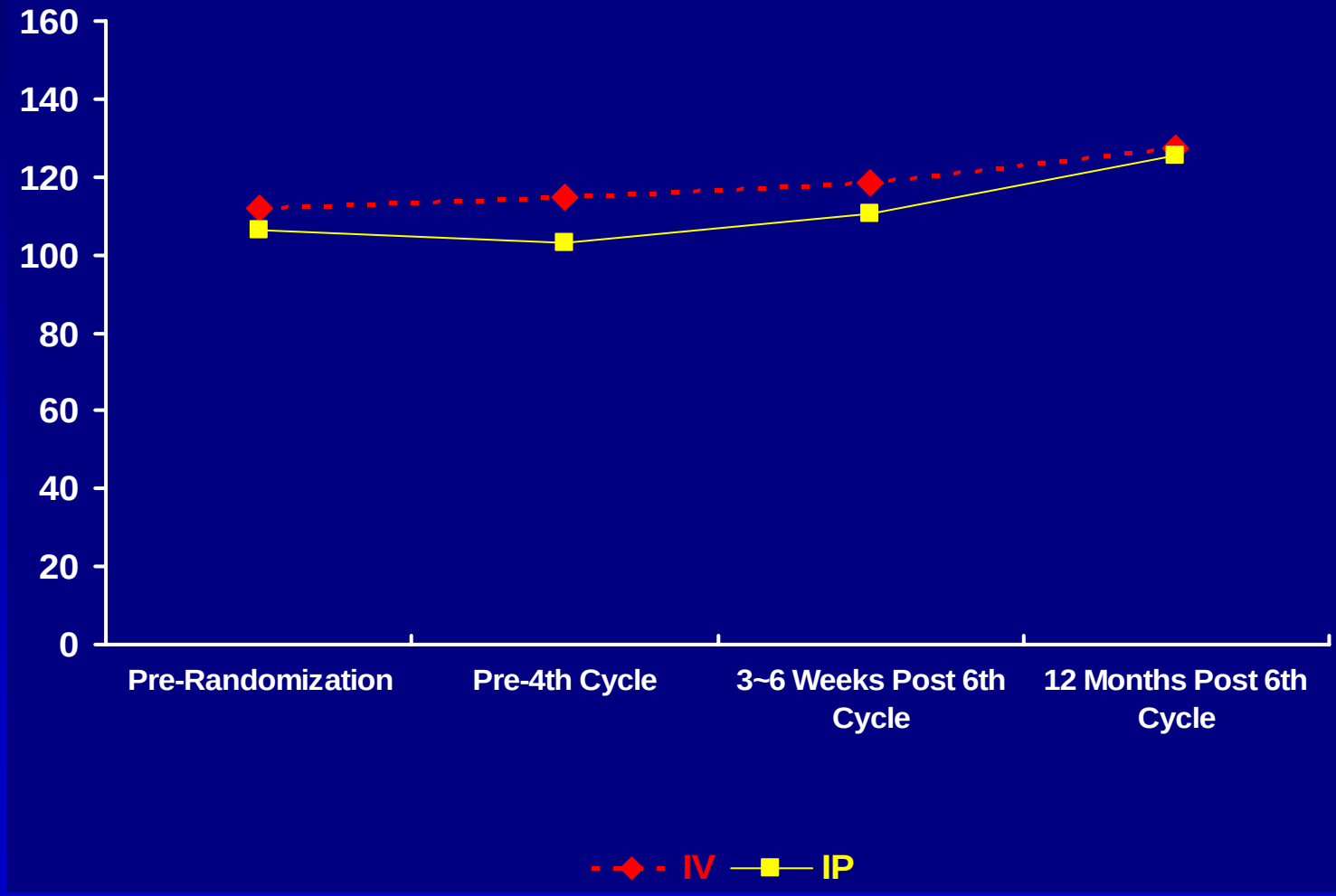
Quality of Life Overview

- Purpose: To describe the QOL differences between IV and IP study arms
  - FACT-O (FACT-G: 27 items; Ovarian subscale: 12 items)
  - FACT-Trial Outcome Index (PWB,FWB,Ov)
  - FACT-GOG/NTX: 11 items (Huang et al, 2006)
  - FACT-GOG/Abd Discomfort: 4 items (Wenzel et al, 2004)

Results – FACT-O

- QOL was significantly worse in the IP group before cycle 4 and 3-6 weeks after treatment ($P<0.01$)
- No significant QOL differences at one year

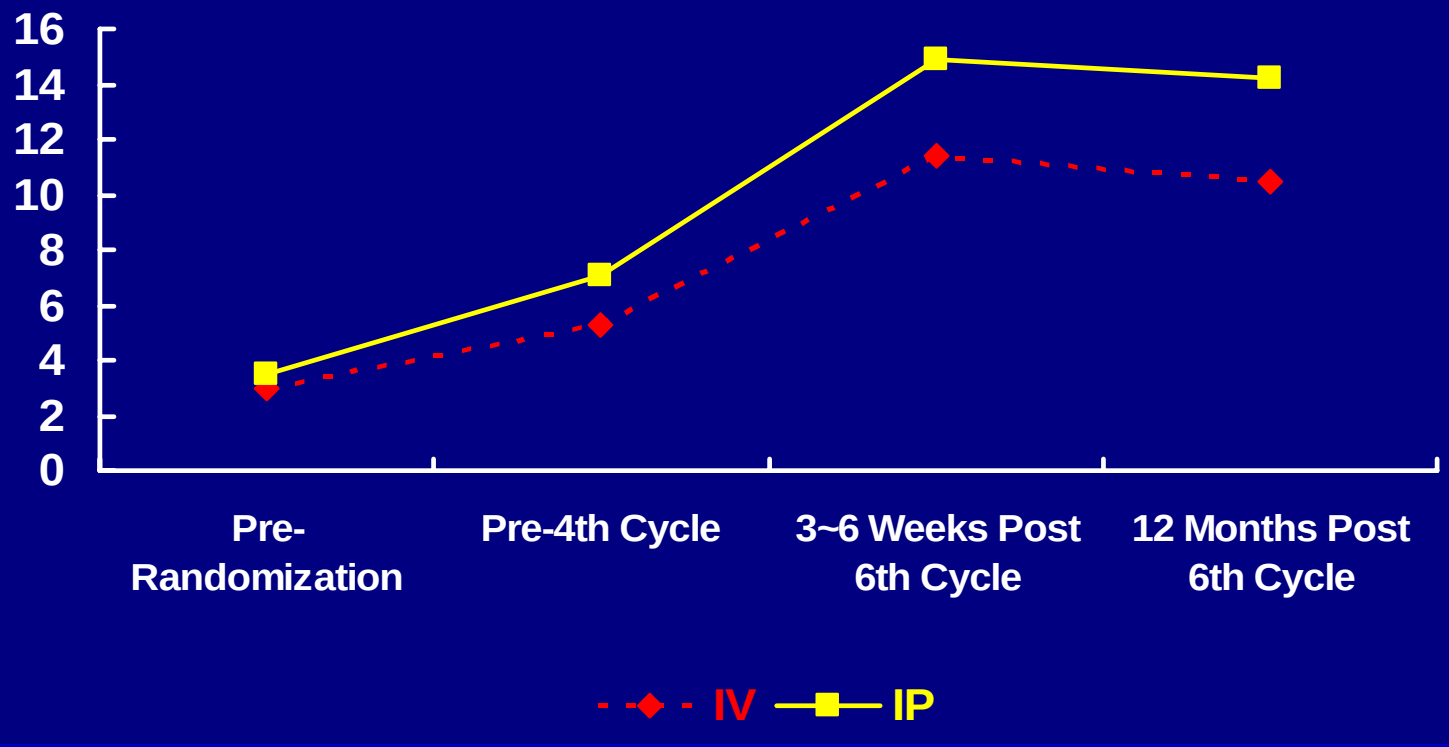
Patient-Reported FACT-O Scores



Results - Neurotoxicity

- Neurotoxicity was significantly worse in the IP arm 3-6 weeks after completing chemotherapy (P=0.0004)
- Neurotoxicity was significantly worse in the IP arm one year later (P=0.0018)

Patient-Reported Neurotoxicity Scores



Patient-Reported Neurotoxicity as an End Point

- **Self-assessment tool enables patients to score their experience with platinum paclitaxel related peripheral neuropathy**
  - 11-item subscale: sensory, motor, hearing and dysfunction for full range of sensory and functional concerns
  - 4-item sensory neuropathy items efficient separation of groups re: chemo-induced NTX
    (Huang et al, 2006. GOG177)

Neurotoxicity: PRO Example

Huang et al, 2006

- 11-item Subscale reliably and validly assesses platinum/paclitaxel-induced neurological symptoms
- 4 sensory items accounted for 80% of treatment differences and 63% of change in NTX scores
  - I have numbness or tingling in my hands/feet
  - I feel discomfort in my hands/feet

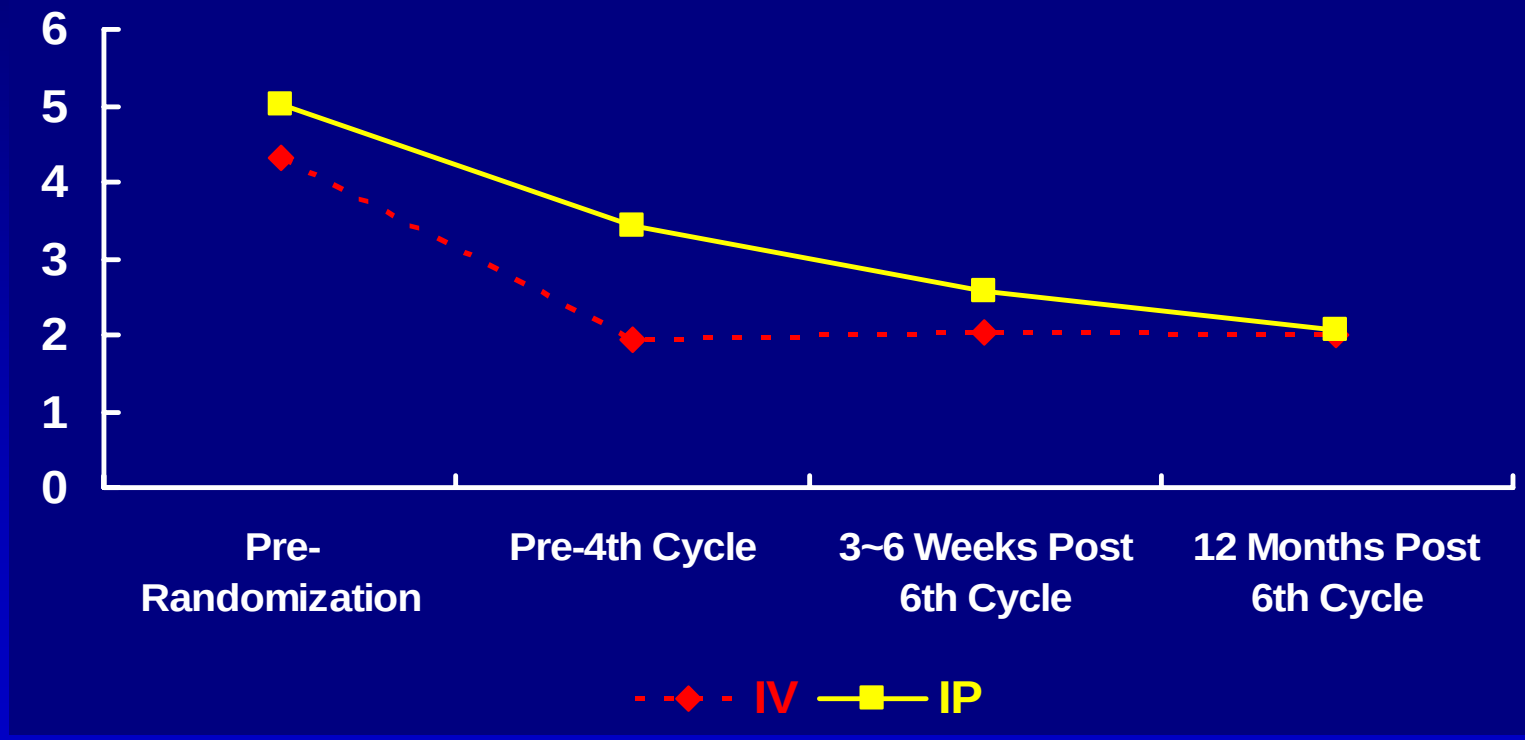
Neurotoxicity: PRO Example

- 11-item scale: internal reliability, construct validity, criterion validity, sensitivity to treatment differences, responsiveness to treatment cycles
- 4-item scale: efficient way to differentiate groups, but misses motor or functional problems
  - Requires further validation

Results – Abdominal Discomfort

- Abdominal Discomfort was significantly worse in the IP arm prior to cycle 4 ($P<0.0001$)

Patient-Reported Abdominal Discomfort Scores



Abdominal Discomfort

- **I have pain.**
- **I have cramps in my stomach area.**
- **I have pain in my stomach area.**
- **Stomach pain interferes with my daily functioning.**

Abdominal Discomfort: PRO Example

Among 205 women on the IP therapy arm, 138 completed the AD subscale prior to cycle 4.

Internal consistency subscale alpha=0.87.

Average inter-item correlations = 0.61.

Item-scale correlations, corrected for overlap, ranged from 0.63- 0.89.

- Item correlation with other scales (including FACT-G (PWB, SWB, EWB, FWB), FACT-O Subscale, and FACT/GOG-NTX Subscale) ranged from 0.01-0.44.

- Both treatments improved abdominal discomfort from baseline to pre-cycle 4 ($p<.01$), effect size (ES) = .56.

- The difference in improvement between study arms favored IV therapy by a margin of 0.9 units (6% of scale; ES=.25).

PRO: Abdominal Discomfort

- The FACT/GOG-AD subscale is a valid and reliable instrument to measure abdominal discomfort.
- The AD subscale is responsive to change over time.
- The AD subscale is a useful tool to document short and long-term effects of abdominal discomfort on QOL.

Conclusions

- Pts who received higher dose IP therapy, compared to those with conventional dose IV therapy experienced
  - More QOL disruption
  - More abdominal discomfort
  - More neurotoxicity
  - HOWEVER, better recurrence-free and OS

Conclusions

- From Baseline to 12 months after treatment
  - Overall QOL improved in both groups
  - Attributed to physical, functional and ovarian-specific subscale improvements
  - Abdominal discomfort improved in both groups from pre-randomization to pre-4th cycle
  - Neurotoxicity worse over time in both groups, especially IP

Implications

- HRQL patient-reported outcomes useful in interpreting treatment implications
- Global HRQL and symptom-specific indices add critical information to IV-IP comparisons

Future Implications

- Continued QOL evaluation critical to
  - Weigh considerable treatment benefits and toxicites
  - Assist in establishing guidelines and safety standards to buffer untoward effects

Future Directions: Use of Patient-Reported Outcomes for Approval of Ovarian Cancer Drugs

- Most likely for recurrent/refractory disease, more symptoms present
- Currently have well-validated health-related quality of life questionnaires, FDA has been reluctant to include these broad concepts in labeling
- Do not have validated symptom index for ovarian cancer

Issues in Developing Ovarian Cancer Symptom Index

- Relevant symptom targets
  - Single index combining multiple symptoms vs multiple measures
  - Disease symptoms vs treatment side effects
- Reliability and validity
- Defining meaningful change

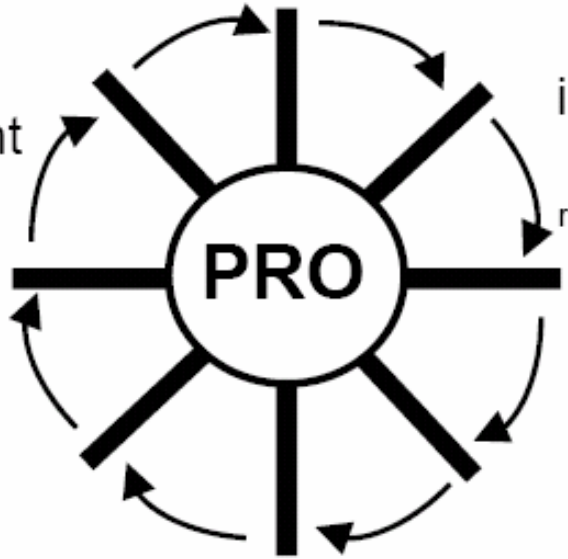
FDA Guidance: Patient Reported Outcome Measures

i. Identify Concepts and Develop Conceptual Framework

Identify concepts and domains that are important to patients.
Determine intended population and research application.
Hypothesize expected relationships among concepts.

ii. Create Instrument

Generate items.
Choose administration method, recall period, and response scales.
Draft instructions.
Format instrument.
Draft procedures for scoring and administration. Pilot test draft instrument. Refine instrument and procedures.



iii. Assess Measurement Properties

Assess score reliability, validity, and ability to detect change.
Evaluate administrative and respondent burden. Add, delete, or revise items.
Identify meaningful differences in scores. Finalize instrument formats, scoring, procedures, and training materials.

iv. Modify Instrument

Change concepts measured, populations studied, research application, instrumentation, or method of administration.

Symptoms in Recurrent Ovarian Cancer

- No single cardinal symptom that is expected to improve with treatment
- Measure(s) need to target multiple symptoms
- When possible, symptoms of disease (expected to improve) should be measured separately from treatment side effects (expected to worsen)

Symptoms in Women with Recurrent Ovarian Cancer

(See, Basen-Engquist, Kavanagh, et al 2005)

- Interviewed 50 women with platinum resistant recurrent ovarian cancer
- About to receive chemotherapy or hormonal therapy
- Completed the FACT-O and a symptom checklist with symptoms from Memorial Symptom Assessment Scale, EORTC QLQ-OV28
- For next stage of index development
  - Identified most frequent, severe, and distressing symptoms
  - Added symptoms from patients below the 25th percentile in QOL
  - Evaluated overlap between items

Symptoms Identified by Patients

- Lack of energy
- Worrying
- Pain
- Aches/pain in muscles and joints
- Feeling bloated
- Feeling nervous
- Difficulty sleeping
- Feeling sad
- Problems with sexual interest/activity

overlap

- Feeling full too quickly when eating
- Sweats
- Hot flushes
- Constipation
- Clothes feel too tight
- Feeling drowsy
- Passing gas
- Heartburn
- Irritable
- Diarrhea

High severity · High distress · High frequency · High severity, low QOL patients

Symptoms Identified by Health Care Providers (Cella, Paul, et al. 2003)

- Most Important Symptom Targets, Nurses/Physicians
  - **Fatigue**
  - Vomiting
  - **Pain**
  - Nausea
  - **Stomach swelling**
  - **Worry condition will get worse**
  - **Content with present QOL**
  - Stomach cramps

Issues in Developing Symptom Index

- One composite index (can ‘symptoms’ be considered a single domain?) or multiple measures of individual symptoms (e.g., fatigue, pain, GI symptoms, psychological symptoms)
- Measures exist for fatigue, pain, psychological symptoms, but not GI symptoms
- What to do with mood? Does mood disturbance occur as the result of having cancer, or can the tumor and treatment exert a more direct effect on mood? Does mood ever belong in the label?

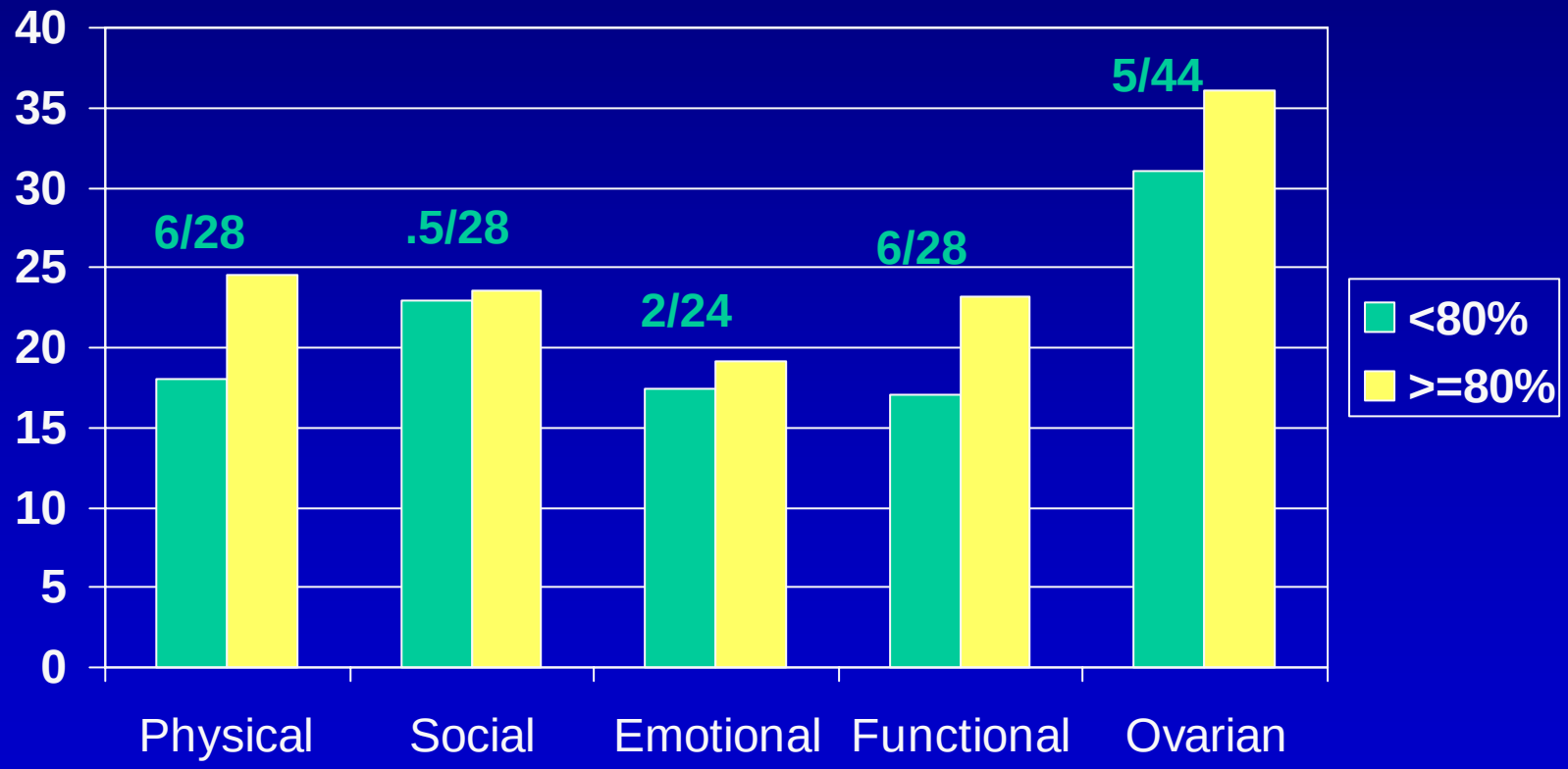
Reliability and Validity

1. Reliability: Extent to which instrument is free of measurement error
   - Internal consistency
   - Test-retest
2. Responsiveness: Ability of instrument to detect change over time
3. Validity: Whether the instrument is measuring what it intends to measure
   - Face validity
   - Construct validity
4. Meaningful differences

Minimum Important Difference

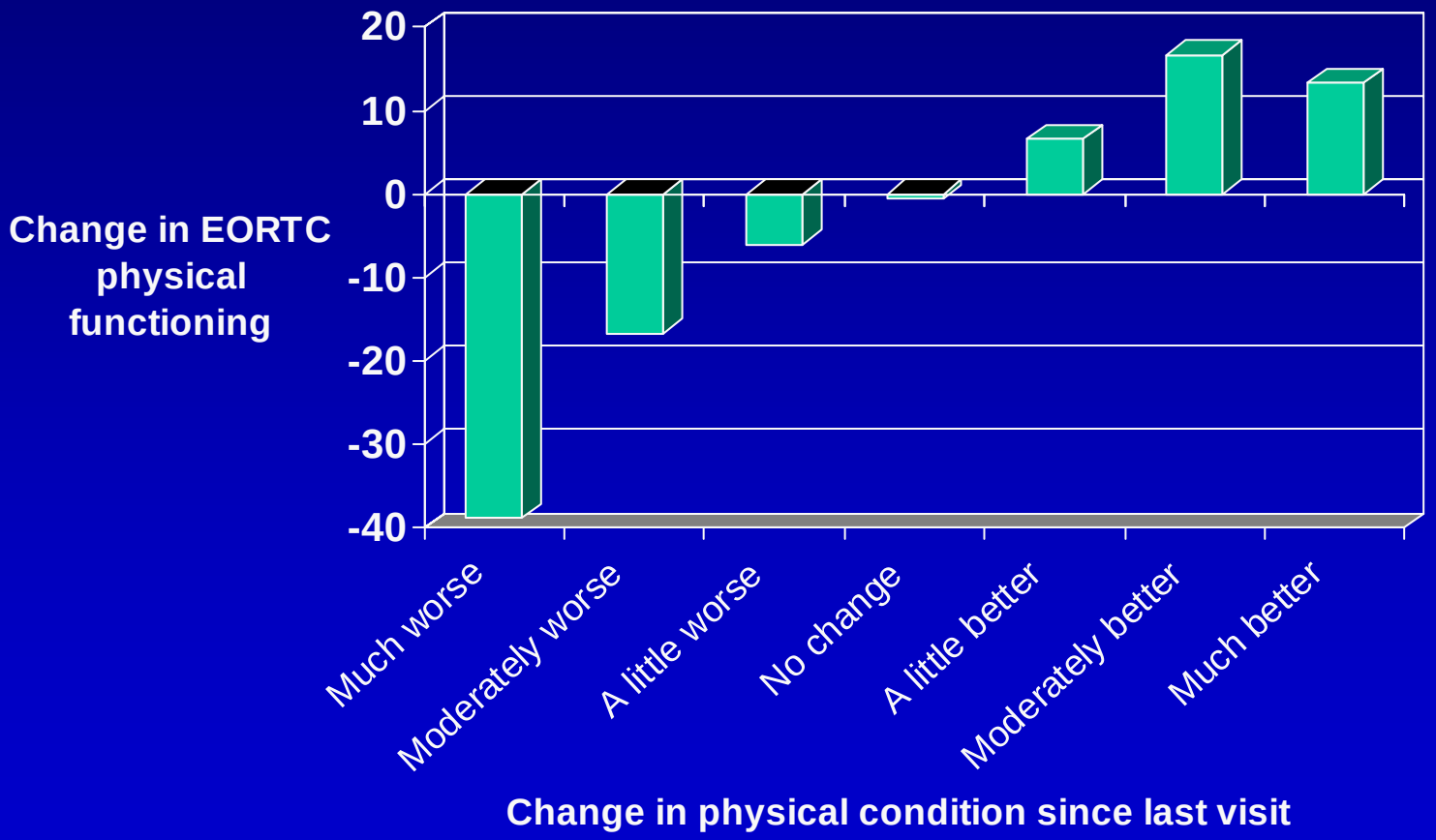
- Anchor-based
  - Anchor to another patient reported outcome, or patient rating of change
  - Anchor to clinical change
- Statistical - Distribution based
  - E.g., .50 standard deviations units = medium effect size (Cohen)
  - 1 Standard Error of Measurement

FACT-O Scores by Extent of Return to Usual Activities



Change in Ovarian Cancer Patients on 2nd- or 3rd-line Chemotherapy

(Doyle, et al, 2001)

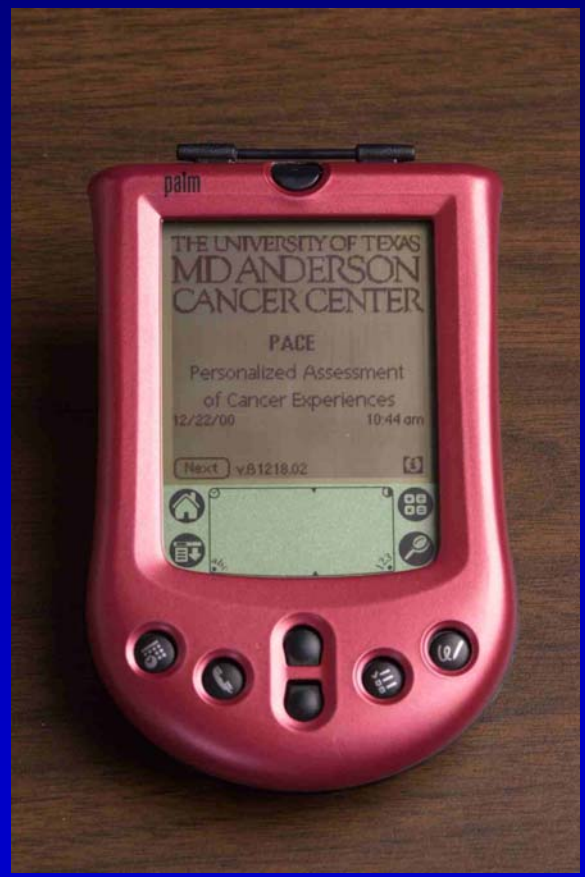


Combination of Anchor-based and Statistical Methods (Cella, Eton et al, 2002)

	Chg for ↑ Hgb	Chg for ↓ Hgb		½ SD
Fatigue	4.9	6.0	6.7	2.7
FACT-G	4.0	6.5	8.8	5.3
FACT-An	11.7	13.1	16.6	6.6
TOI-Fatigue	9.3	10.5	12.0	4.8
TOI-Anemia	11.1	12.0	14.0	5.6

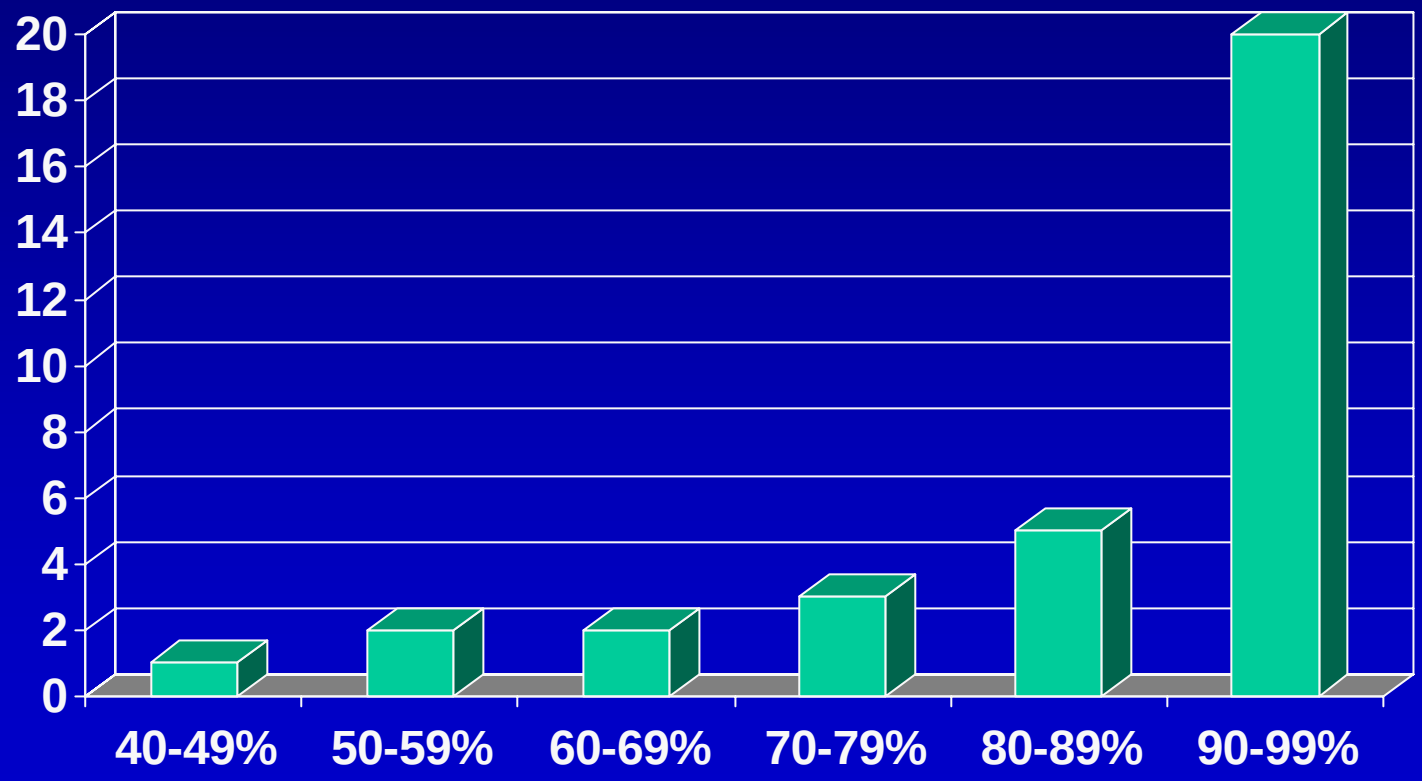
New Directions: Real-time Assessment Methods

- Electronic Diaries
- Patients use handheld computers, interactive voice response systems, etc. to record PRO information
- Advantages
  - Better recall
  - Ecological validity
  - Can identify short term pattern in PROs (e.g., diurnal patterns)
  - Better understanding of within subject variability

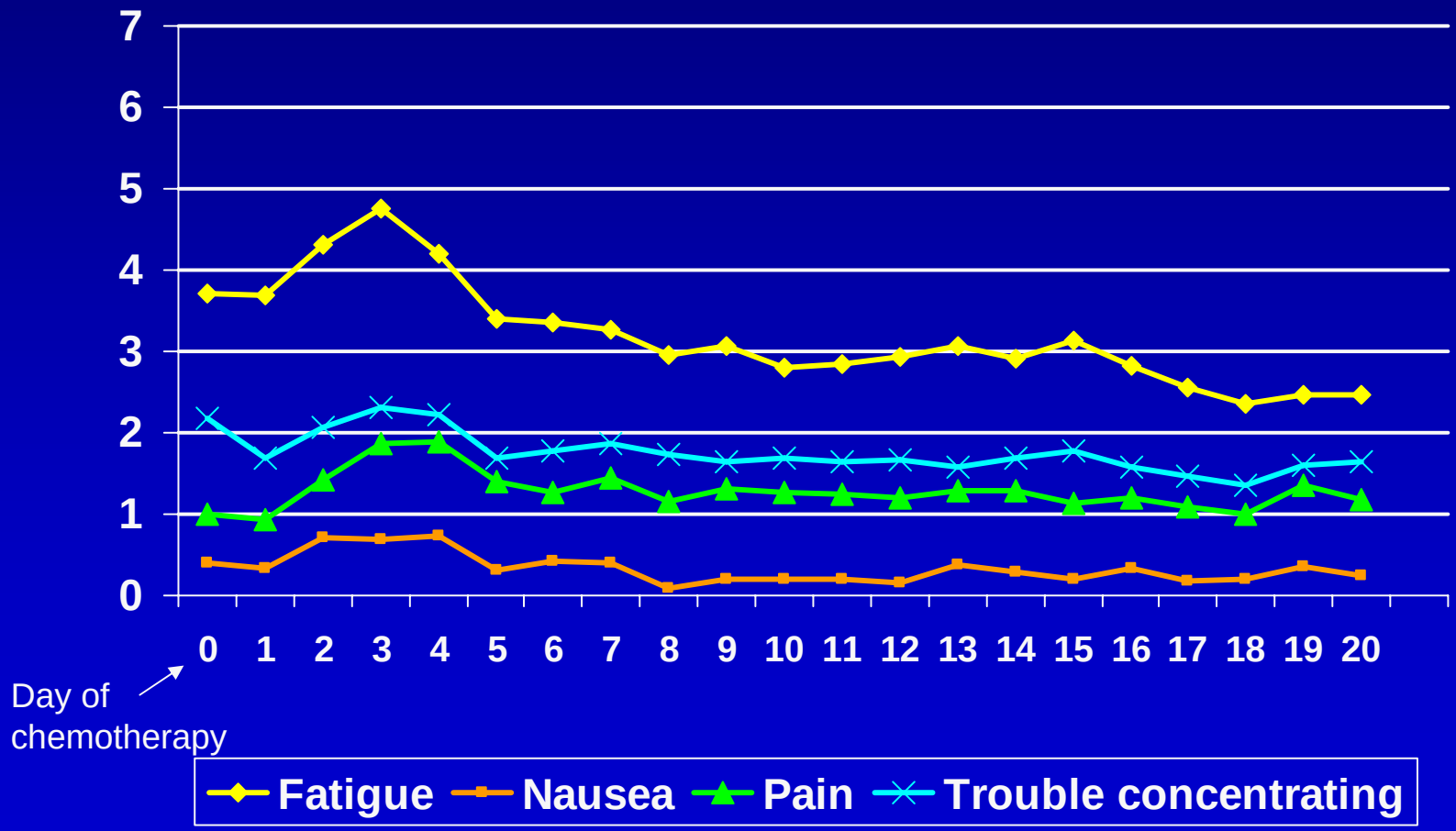


Feasibility: Data Completeness

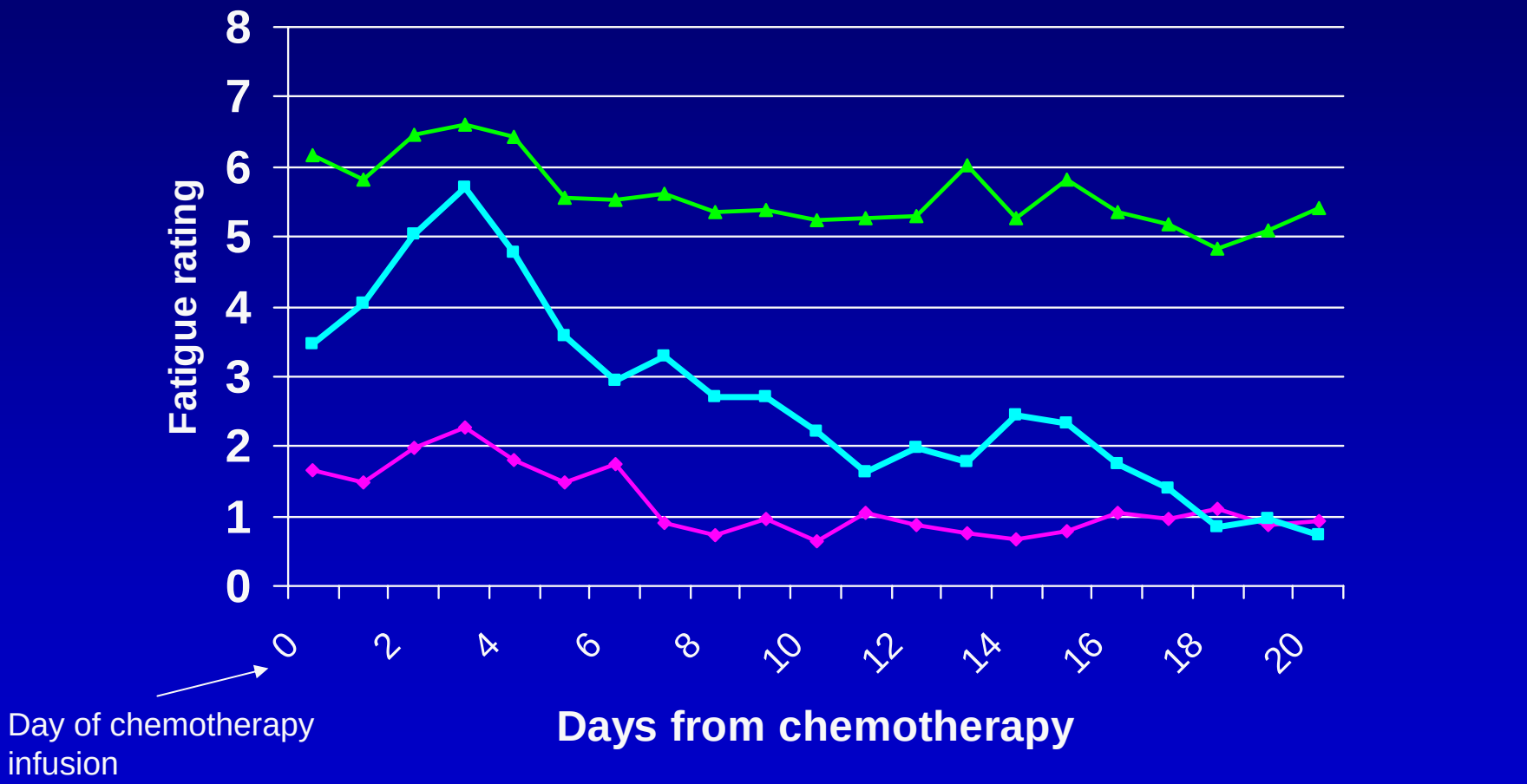
Average % of assessments completed: 86%



Patterns of Fatigue and Other Symptoms over Chemotherapy Cycle



Fatigue Pattern Clusters, 3 Cluster Solution



New Directions: Computer Adaptive Testing

- Computer algorithm uses patient's previous response to select next question – will choose question that will provide maximum information. Results in briefer, more precise measures
- PROMIS Initiative: Evidence-based conceptual framework, common patient endpoints, large and well-tested repository of questions to measure most common and important symptoms and functional concepts

PRO Issues for Further Study

- PROs assist in documentation of symptoms
  - What symptoms are most important to patients?
  - How are PROs related to clinical outcome?

The optimal treatment for ovarian cancer remains uncertain.

"Therefore, HRQL data can provide unique information that leads to the choice of effective treatments, rejection of ineffective interventions, and clarifies the tradeoffs between management strategies …" (Mayo Clinic Proceedings, 2003)