Phase III Trials: Case Studies and Lessons Learned

Patricia A. Ganz, M.D.

Carolyn C. Gotay, Ph.D.

Overview of Presentation

- Before Lunch
  - Introduction and methods
  - Challenging trials and their outcomes
  - Successful trials and their contributions

After Lunch

- Synthesis of “Lessons Learned”
- Roundtable discussion and audience response

Introductory thoughts….

- Patient-reported outcomes have a long history in phase III cancer treatment trials
  - Sugarbaker & Barofsky: Sarcoma limb-sparing surgery, use of existing rehabilitation scales (Surgery, 1982)
  - Priestman & Baum: Advanced breast cancer, use of cancer specific LASA scales (Eur J Ca, 1980)
- Initial challenges and barriers: staff resistance, inadequate measures, concerns about burden and costs
- Some of these same challenges persist today

Methods

- Brief survey questionnaire sent to leaders of PRO activities in US cooperative groups and CCOP bases (May 2006)
  - Nominate up to 3 trials that were successful or unsuccessful.
  - Why did you nominate this study? Please tell us in three sentences or less.
  - Please list publications based on this trial, if any. Not necessary to have publications for a trial to be considered.
  - Request for protocol and publications for those selected for discussion.

Results

- Response from 6 cooperative groups; none from CCOP research bases
- Variety of examples, with many from the groups with mature PRO efforts
- Presentation today will focus on completed studies
- Selections made from among 20 examples suggested

Challenging trials

- S9509: Advanced lung cancer
- CALGB 9481: Colorectal cancer & hepatic metastases
- NSABP B-23: Adjuvant breast cancer
- S9208: Early stage Hodgkin’s Disease

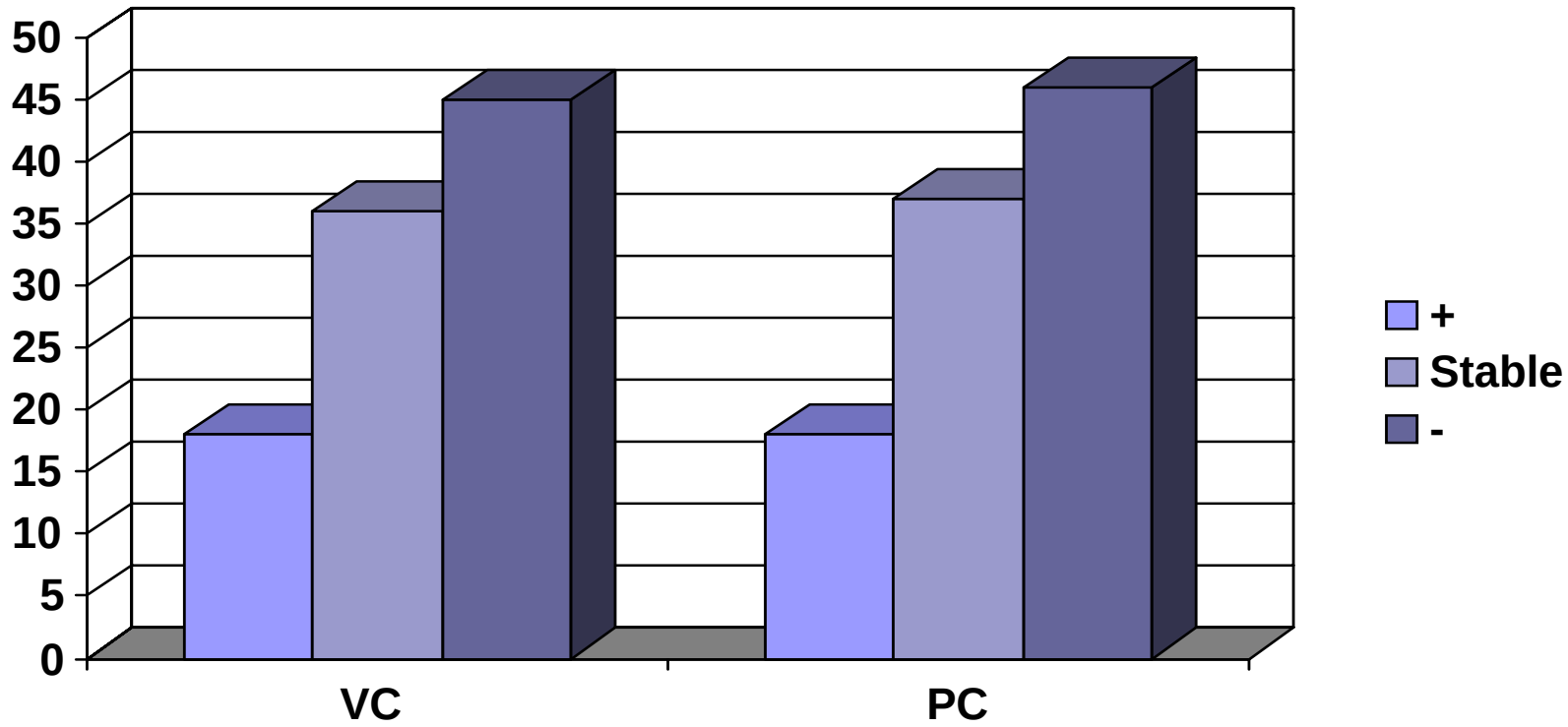
S9509: *PC vs. VC in Advanced Non-small Cell Lung Cancer*

- N=408, 222 in PRO study
- PRO measure: FACT-L
  - Baseline, 13 and 25 weeks
- Primary outcome: survival
- Other measures: costs, toxicity

S9509 PRO Question and Study Findings

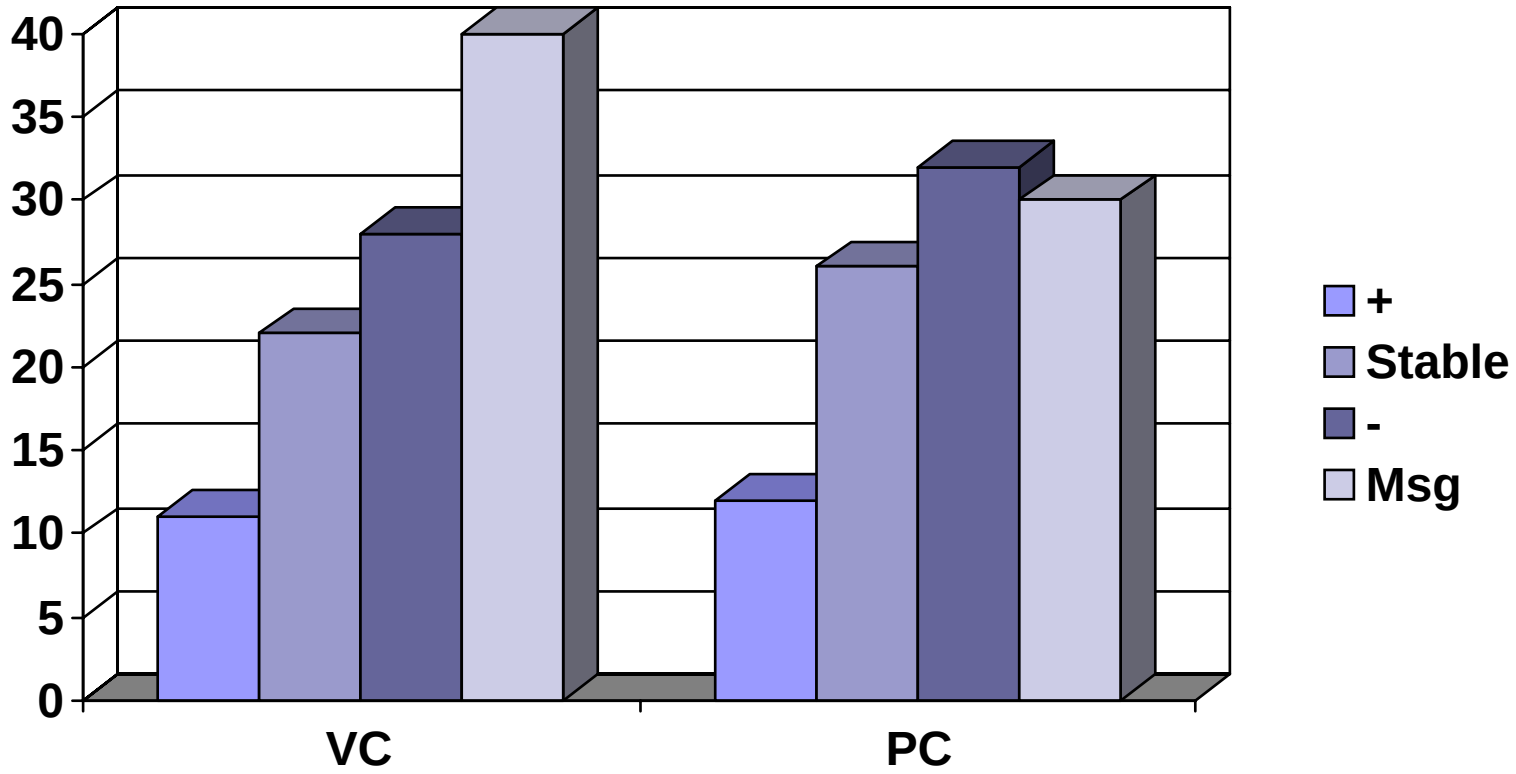
- PRO question: “Secondary objectives were to compare toxicity, tolerability, QOL, and resource utilization between the two arms.”
- Findings: no differences in survival or QOL by arm, some toxicity and cost differences

S9509 QOL Results: Patients with questionnaires



Data based on 13 week endpoint; y-axis is % in each category

S9509 QOL Results: All patients



Data based on 13 week endpoint; y-axis is % in each category

S9509: Challenges

- QOL assessment added halfway through study
- Timing of assessments
- Missing data
  - Completion rates: 91%, 68%, 47%
- Interpretation of “stable QOL” status
- Publication in different journals

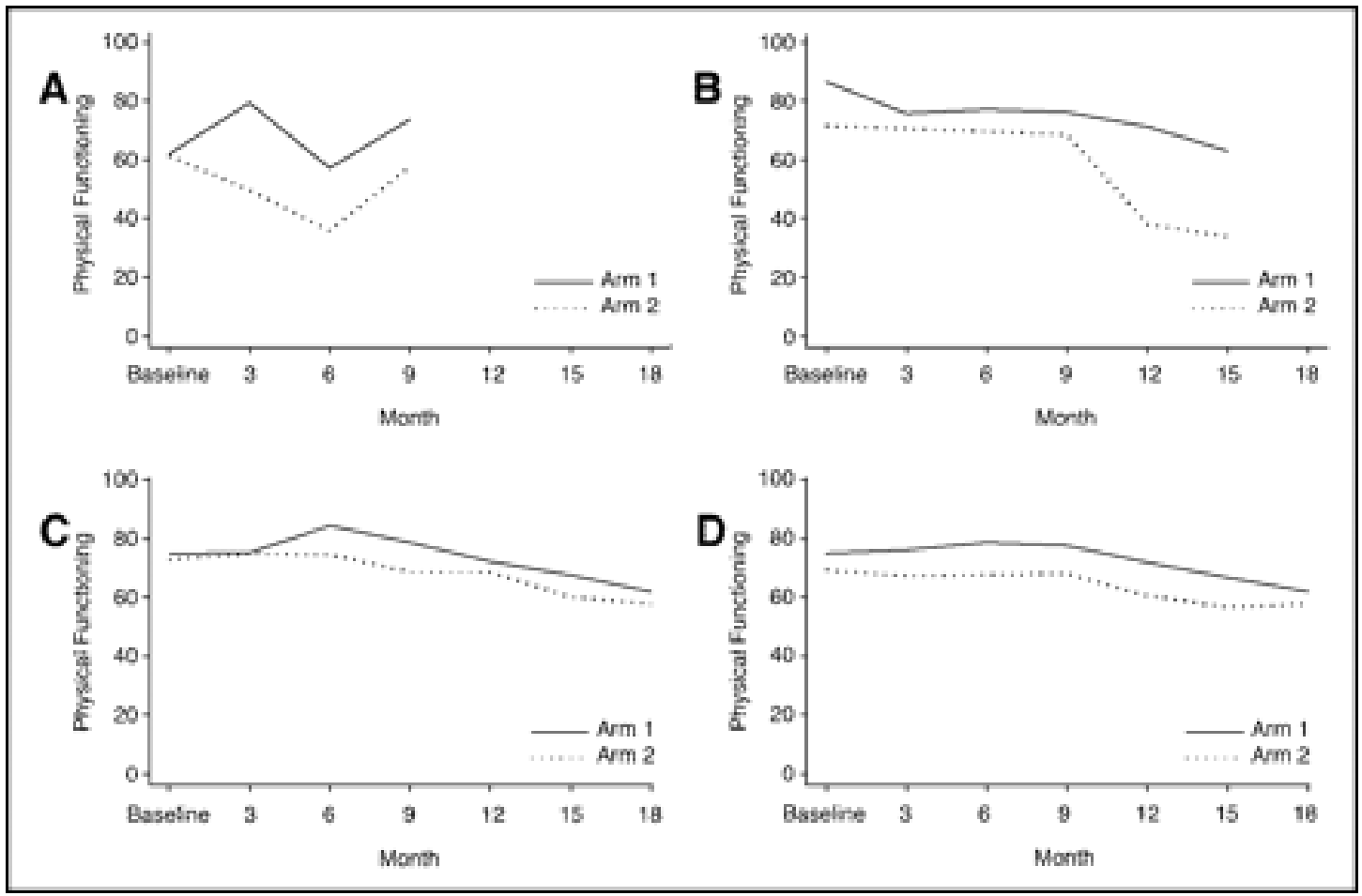
CALGB 9481: *Hepatic Arterial Infusion (HAI) vs. Standard CT in Colorectal Cancer with Unresectable Liver Metastases*

- N=135
- PRO measures: SF36 (physical, role, social, general health perceptions), Memorial Symptom Assessment Scale, MOS social support, MOS sexual functioning
  - Baseline, every 3 months for 18 months
- Primary outcome: survival
- Other measures: toxicity, cost, biomarkers

CALGB 9481: PRO Question and Study Findings

- PRO question: “Secondary end points were tumor response, toxicity, quality of life (QoL), and cost effectiveness.”

CALGB 9481:QOL Results



CALGB 9481: Challenges

- Missing data
  - 47% completed all assessments (n=7)
  - Unclear degree of missing data at primary endpoints of 3 and 6 months
- Multiple PRO measures
  - No information re: symptom and MOS measures
  - Treatment differences were predicted for all 4 SF-36 scales, but only physical functioning showed an effect

NSABP B-23: *Adjuvant AC vs. CMF in Node Negative, ER- Breast Ca*

- Design: RCT with comparison of 4 cycles vs. 6 cycles of treatment, different schedules (q3wk vs. q4wk)
- PRO Instruments: FACT-B, SF-36 vitality scale, health rating scale, symptoms
- Key PRO question: if treatment outcomes are similar, will PROs be better with one regimen or the other

NSABP B-23: PRO implementation challenges

- Treatment trial opened to accrual in 1991; QOL substudy opened in 1997 with only 18 month accrual in selected sites
- Treatment trial closed in 1998, as did QOL component
- Target accrual for the QOL study was 200 and only 160 enrolled; only 69 of 111 institutions participated and not all eligible enrolled in QOL study

NSABP B-23: Complex data collection schedule; missing data

Table 1. Chemotherapy and QOL assessment schedule[a]

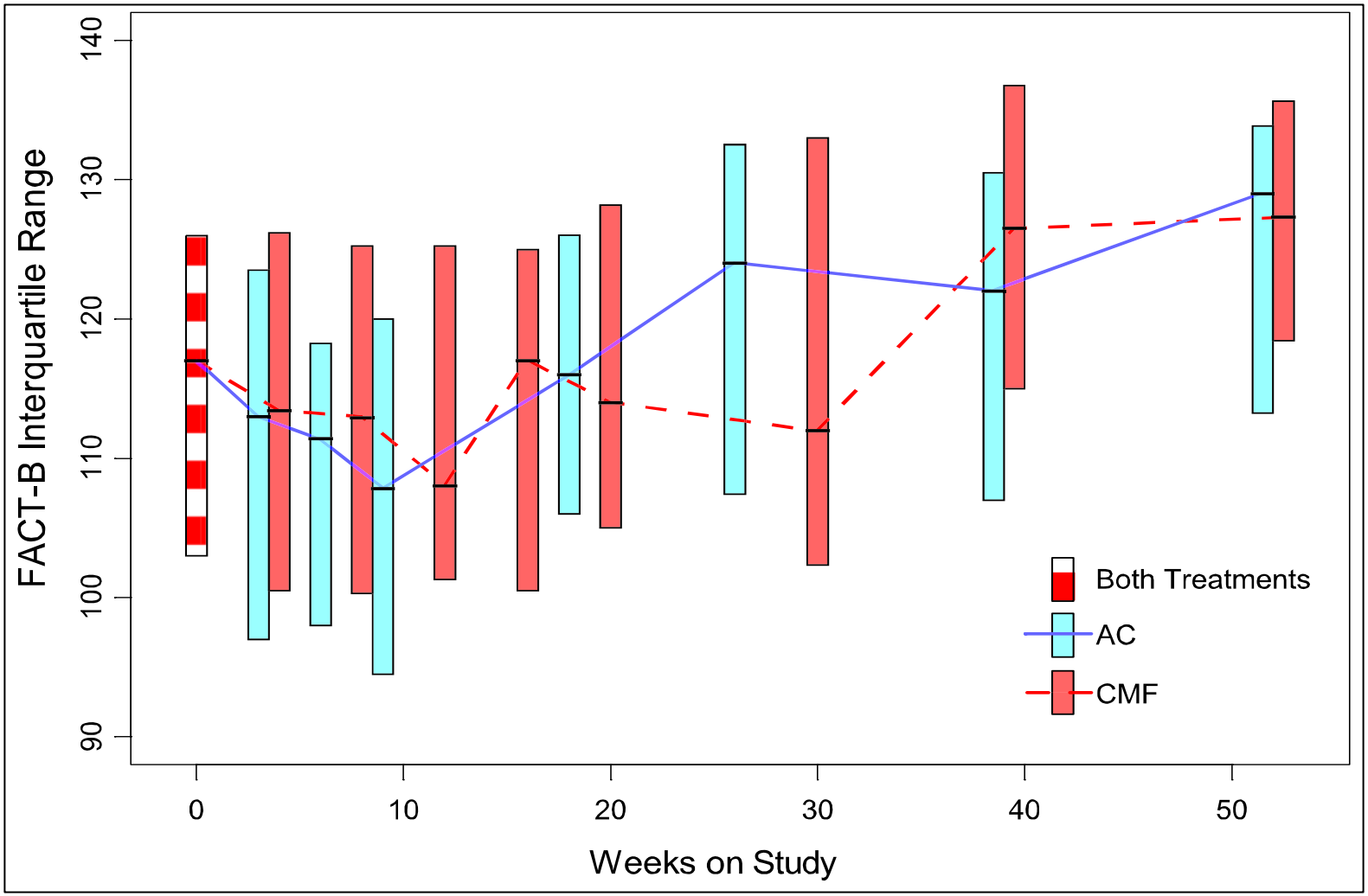
		Baseline	Week after randomization											
			3	4	6	8	9	12	16	18	20	26	30	39	52
AC	Start of cycle	1	2		3		4								
	QOL assessment	√	√		√		√			√		√		√	√
	# Expected forms	81	81		81		79			79		79		78	78
	% Submitted	96	70		80		77			62		59		50	54
CMF	Start of cycle	1		2		3		4	5		6				
	QOL assessment	√		√[b]		√[b]		√[b]	√[b]		√[b]		√[c]	√	√
	# Expected forms	79		79		78		78	77		77		31	74	73
	% Submitted	92		82		77		72	71		71		55	54	58

Br Ca Res Trt 86:153, 2004

B-23 QOL Study – Results

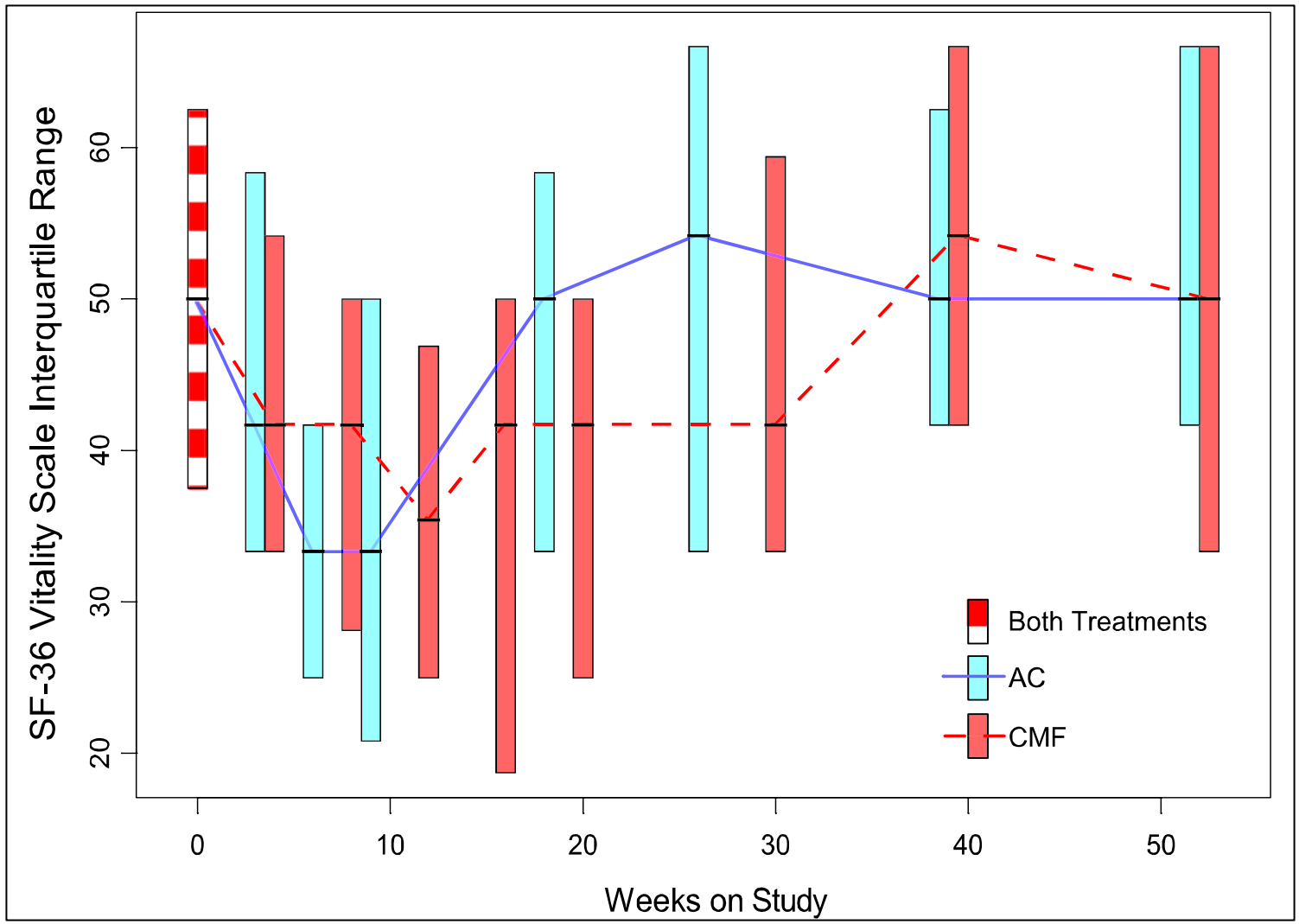
- Overall QOL (FACT-B) was no different between the two treatment arms during treatment, or at 9 and 12 months
- Pattern of fatigue differed between the two treatment arms
- Different pattern of symptoms between the two treatment arms

Cancer-specific QOL in the Year after Randomization to AC or CMF



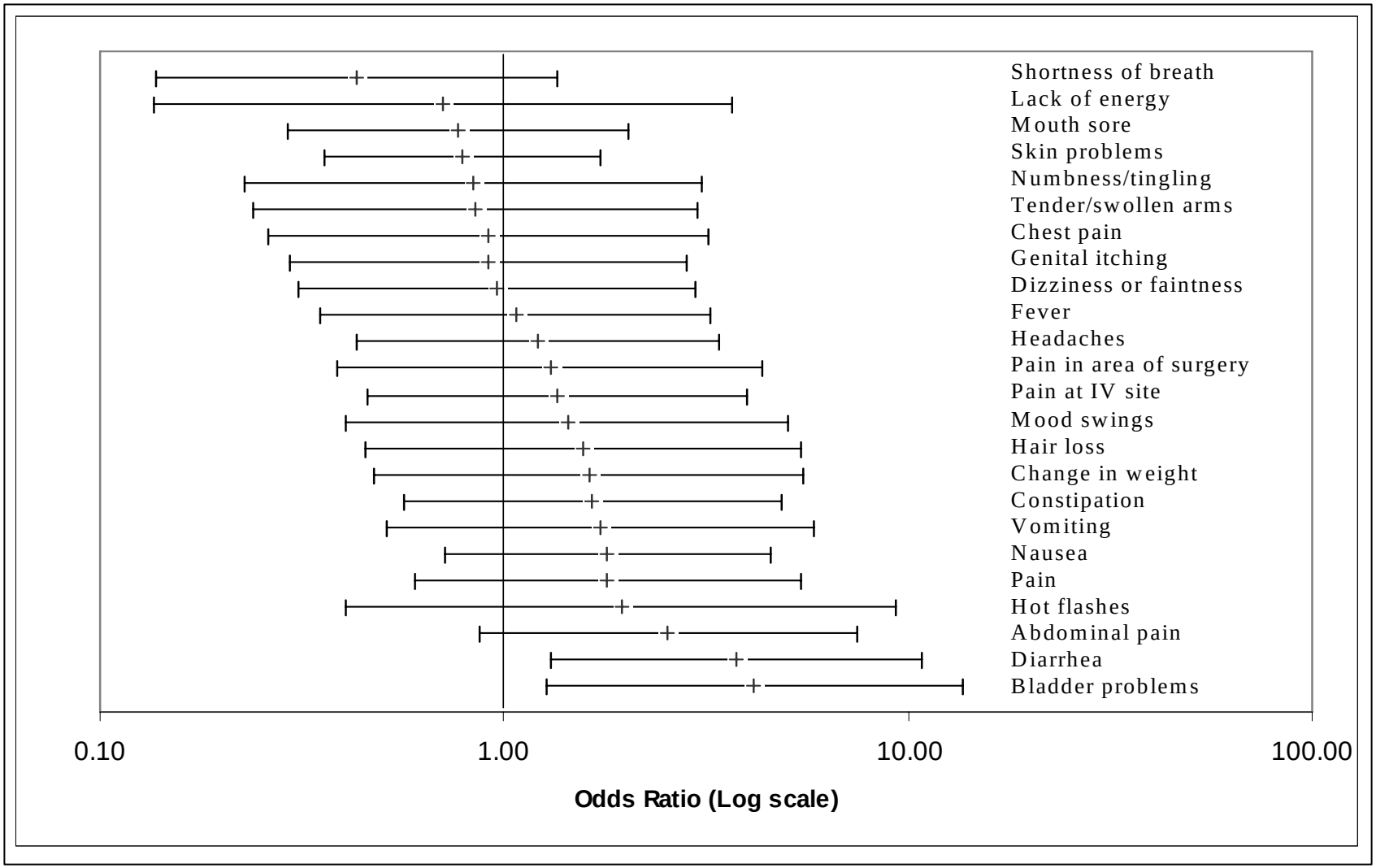
Br Ca Res Trt 86:153, 2004

Energy/Fatigue after Randomization to AC or CMF



B-23: Self-reported Symptoms and Treatment

AC Worse CMF Worse



B-23 QOL – Challenges

- Missing data; too many assessments
- Different timing of assessments
- Sample size barely adequate due to late initiation of PRO study
- Analysis strategy primarily descriptive with many endpoints

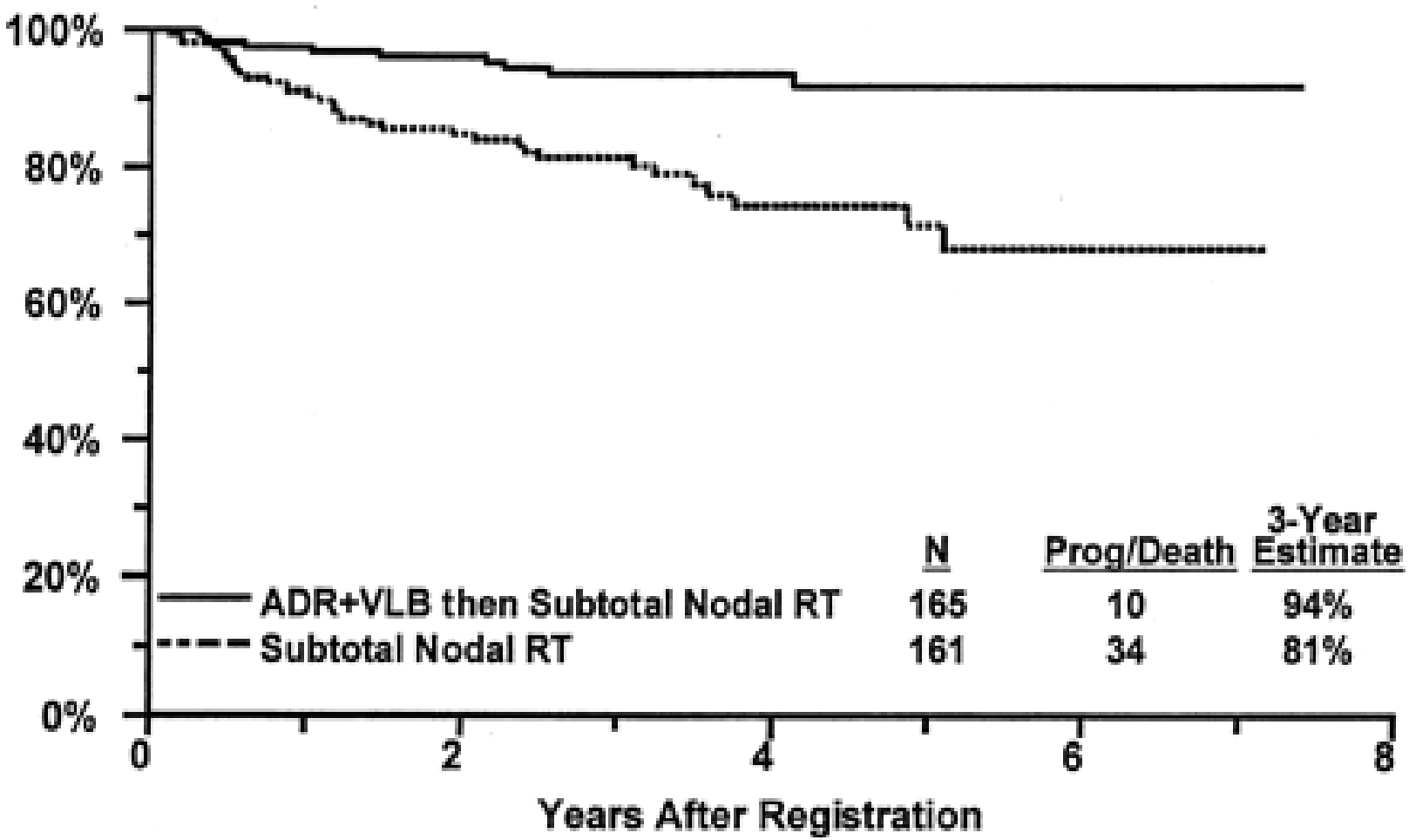
S9208: *Health Status and QOL in Early Stage HD patients treated on S9133*

- Design: RCT with comparison of RT vs. short course chemo + RT
- PRO instruments: CARES-SF; SF-36 Vitality and Health Perception; SDS
- Key PRO Questions: Is short-term morbidity of chemo worth improved DFS? What is the impact of recurrence on survivor QOL?

S9208 Implementation Challenges

- Written as a separate companion trial to S9133 treatment trial
- Started accrual 19 months after S9133 opened
- S9133 closed early due to better than expected response; 326 in treatment study
- Sample size for S9208 less than expected; 224 patients in PRO study

Failure-Free Survival in SWOG 9133



Overall Survival of Patients Treated on SWOG 9133

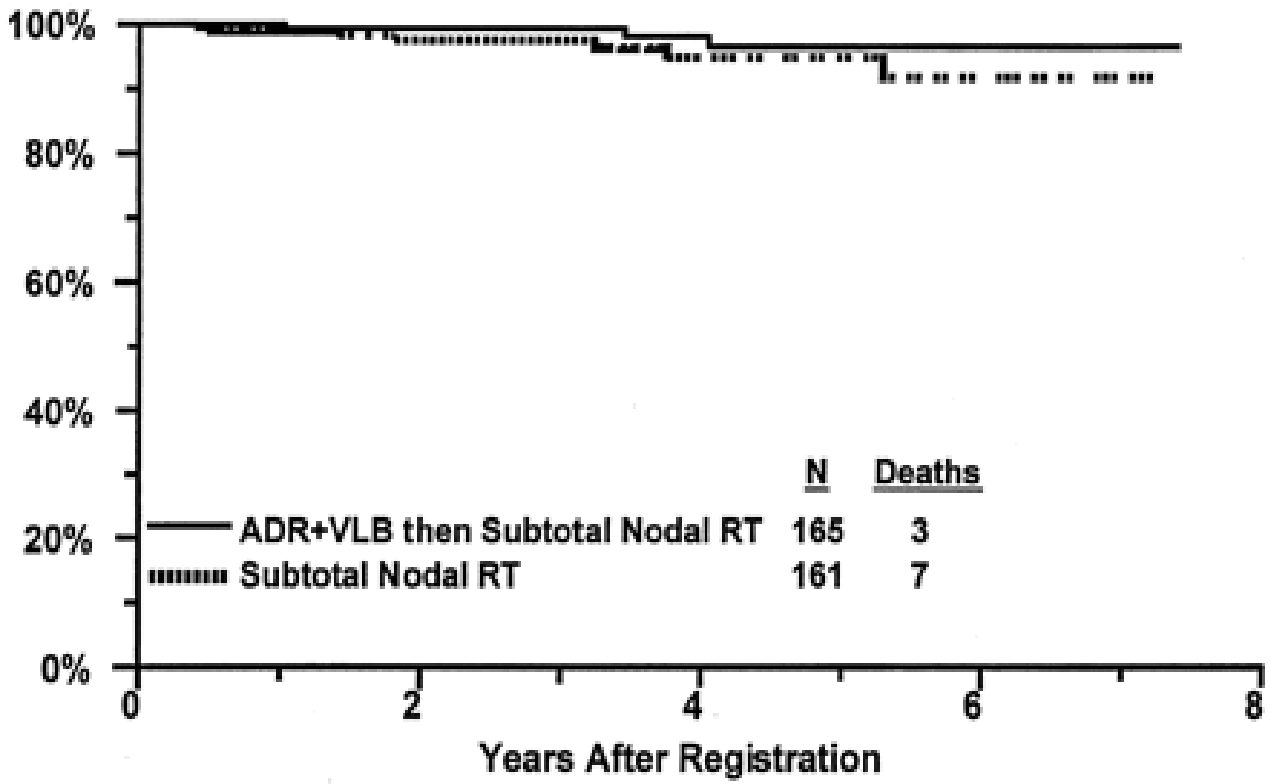


Table 2. CARES-SF Submission Rates by Assessment Time and Treatment Arm: Percentage of Forms Completed for Patients Alive and On-Study, to Specific Assessment Point

	CMT			STLI			Overall
	No. Due	No. Submitted	%	No. Due	No. Submitted	%	%
Baseline							
CARES-SF	124	121	98	120	116	97	97
Symptom Distress Scale	124	121	98	120	118	98	98
6-Month							
CARES-SF	121	100	83	110	87	79	81
Symptom Distress Scale	121	100	83	110	93	85	84
Year 1							
CARES-SF	120	94	78	109	85	78	78
Symptom Distress Scale	120	97	81	109	89	82	81
Year 2							
CARES-SF	120	82	68	107	78	73	70
Symptom Distress Scale	120	84	70	107	77	72	71

Abbreviations: CARES-SF, Cancer Rehabilitation Evaluation System–Short Form; CMT, combined-modality treatment; STLI, subtotal lymphoid irradiation.

High rate of missing data at 1 and 2 years with few deaths or relapses.

JCO 21:3512, 2003

S9208: Other challenges

- Annual PRO assessments required out to 7 years for survivorship endpoints
- No centrally coordinated reminder system; responsibility of study coordinator without financial support
- Continued attrition and lost to follow-up in spite of some financial incentives to sites (CCOP credit and $$)

Successful Trials and Their Contributions

- RTOG 9719: Bone metastases
- S9039: Advanced Prostate Cancer
- CALGB 9221: Myelodysplastic Syndrome
- NSABP B-35: Breast Ductal Carcinoma In Situ

R9719: *Short vs. Long Radiotherapy for Bone Metastases*

- N=898, breast or prostate cancer, moderate to severe pain at up to 3 metastatic sites
- PRO measures: FACT, BPI, HUI III
  - Baseline, 2 and 4 wks, 2,3,6,9,12,18,24, 30,36, 48, 60 mos

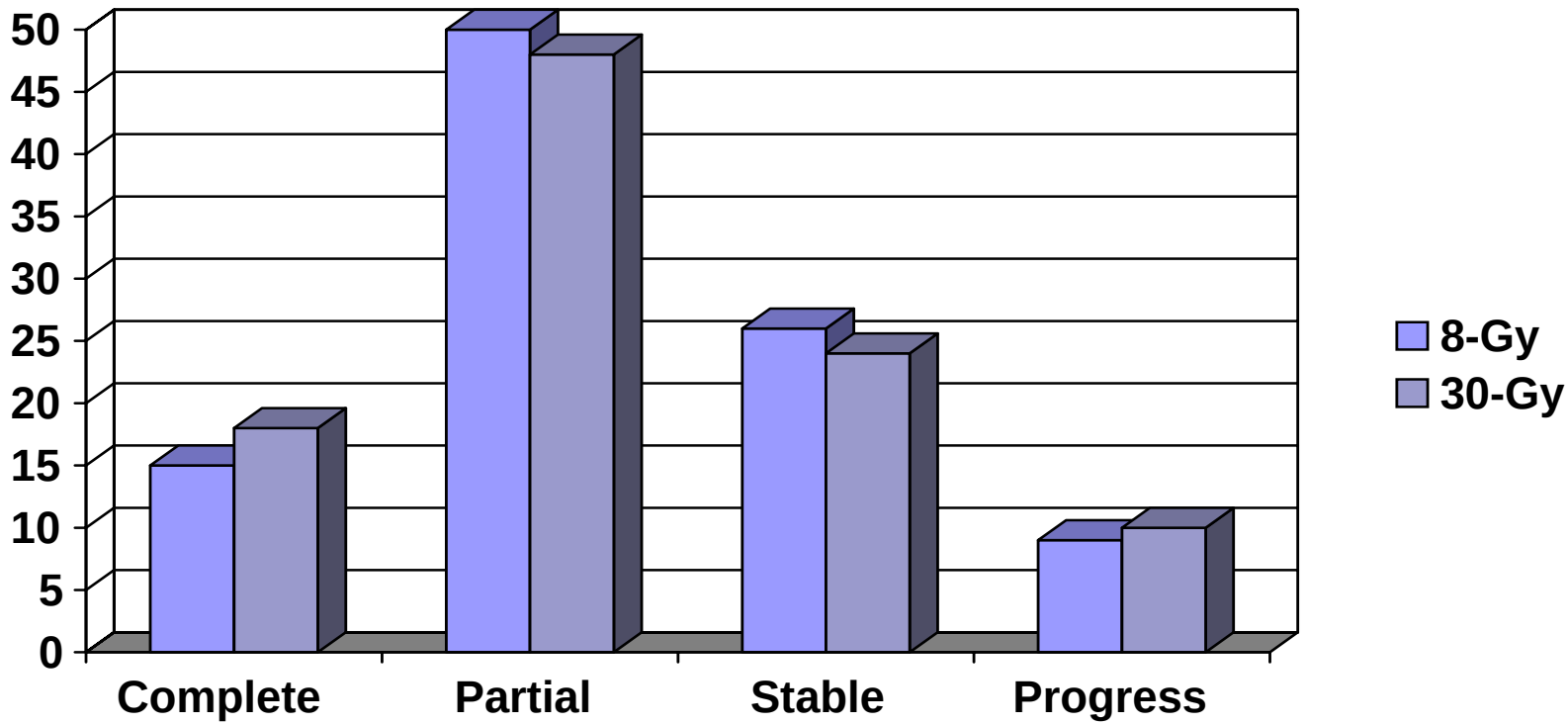
R9719: PRO Question

- Does single fraction XRT (8 Gy) provide equivalent pain and narcotic relief to 10 fractions XRT (30 Gy)?

R9719: Findings

- QOL and HUI will be reported later
- Results based on 3-month data
- Missing data at 3 months: 19% of pts had died or were too ill, 84% of patients who could fill it out did so

R9719: BPI Results re: Pain at 3 Months



R9719: Strengths

- Clear hypothesis, straightforward and readily interpretable outcome
- Important clinical problem with health care cost implications
- Exploration of additional variables: toxicities, fractures,type of analgesic/narcotic, stratification factors
- Possible biological explanation of findings
- Potential for changing practice

S9039: *QOL in Advanced Prostate Cancer Patients Who Received Orchiectomy +/- Flutamide*

- N=739
- PRO measures: Symptom Distress Scale, SF-36 Physical Functioning, Mental Health Index, Social Functioning, SF-20 Role Functioning, symptom items
  - Baseline, 1, 3, 6 months

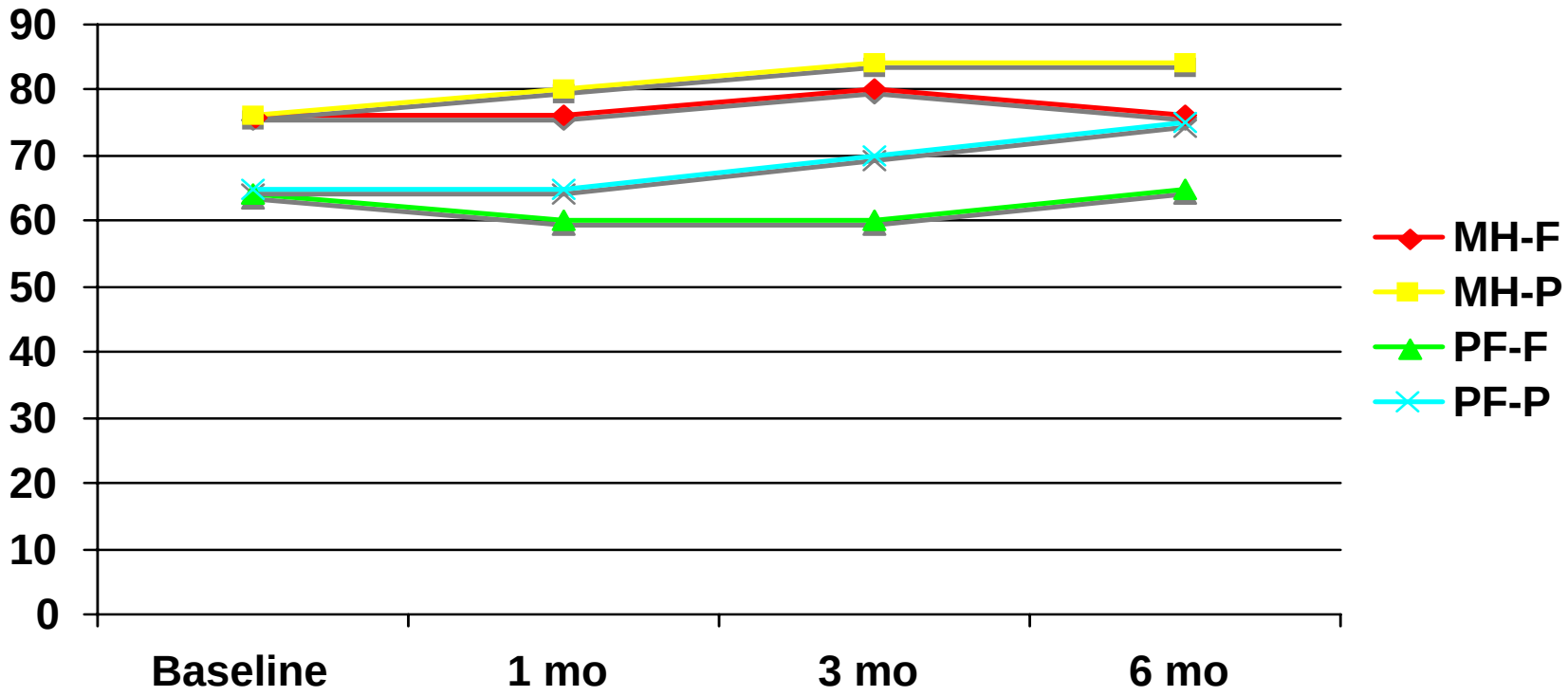
S9039: PRO Question

- To examine PR-QOL during initial 6 moths post-bilateral orciectomy vs. placebo
  - 5 QOL parameters: diarrhea, gas pain, body image, physical functioning, emotional functioning

S9039: Findings

- Companion trial to therapeutic intergroup trial (N=1387), no survival differences
- % questionnaire completion: 98, 88, 86, 81 (baseline, 1, 3, 6 months)

Mental and Physical Functioning Scores Over Time (Md scores)



S9039: Strengths

- One of the first of the “modern era” PRO studies
- Collection of QOL data allowed identification of flutamide treatment effects that would not have been found using only CTC or symptom data
- Selection of measures of particular relevance to this population - interpretation of clinical significance limited
- Amount of missing data low
  - Though advanced disease, patient deaths low during study period
  - Special efforts made by study coordinator
- Possible biological explanation of findings

S9039: Implications

- Collection of QOL data allowed identification of limitation of flutamide treatment that would not have been found using only CTC or symptom data
- Effects extended to both specific symptoms and overall well-being

CALGB 9221: *5-Azacytidine vs. Observation in MDS*

- Design: Phase III RCT with cross-over
- PRO Instruments: EORTC QLQ and MHI administered by telephone interview
- Key PRO Questions: Hypothesis-response to Aza C would result in improved quality of life attributable to better palliation, with less fatigue resulting in improved physical and social functioning and less psychological distress.

JCO 20:2429-2440, 2002

JCO 20:2442-2452, 2002

CALGB 9221: Methods and Results

- Statistics: study had 80% power to detect a medium effect size of 0.57 between treatment arms in three quality-of life measures for the change from baseline to the second follow-up; pattern mixture model analysis to address attrition
- Results: N=191 patients in treatment and PRO study; Aza-C treated patients had improved fatigue, dyspnea, physical functioning and mental health

Response of PRO Measures to Cross-over Therapy

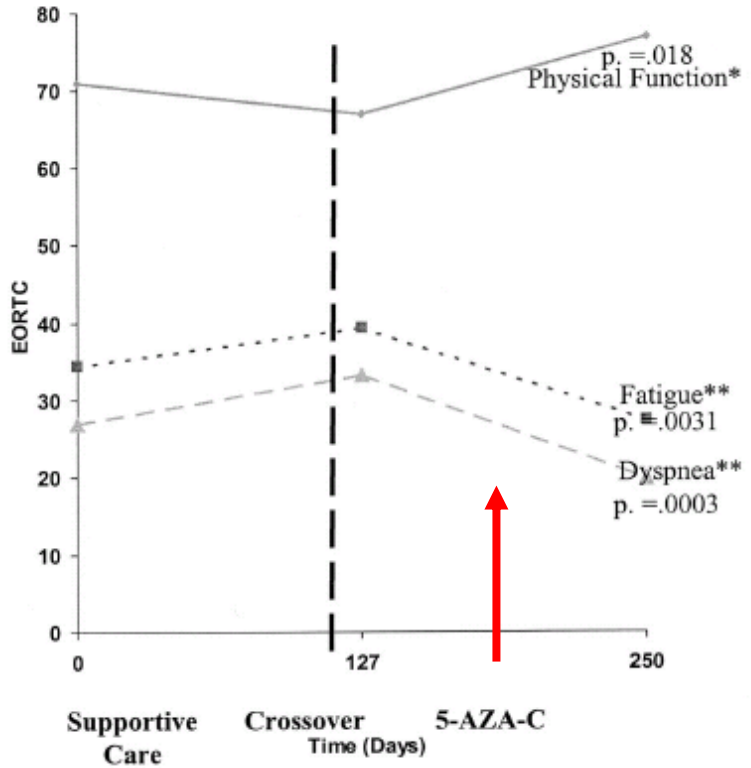


Fig 5. EORTC fatigue, dyspnea, and physical functioning of patients who cross over from supportive care to Aza C (n = 30). *Higher scores indicate better functioning. **Lower scores indicate symptom improvement.

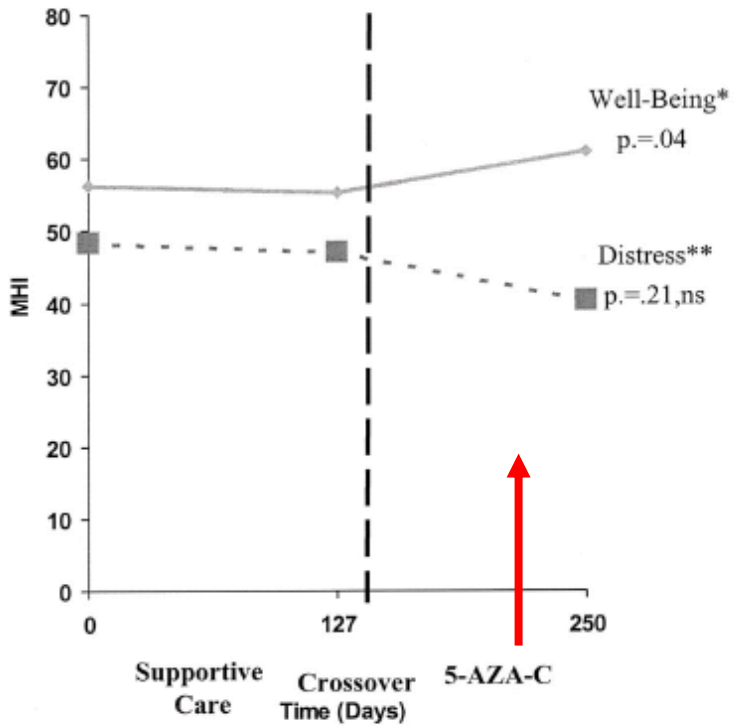


Fig 6. MHI physiological distress and well-being of patients who cross over from supportive care to Aza C (n = 30). *Higher scores indicate better well-being. **Lower scores indicate less distress.

Kornblith, et al.JCO 20:2442-2452, 2002

CALGB 9221: Strengths

- PRO study integrated into trial design and protocol
- PRO outcomes integrated into primary study report and companion detailed paper reported in same journal
- PRO outcomes were instrumental in drug approval process

NSABP B-35: *RCT Comparing Anastrozole with Tamoxifen in Postmenopausal Patients*

with DCIS Undergoing Lumpectomy with RT

- Design: Double blind, placebo controlled trial with PRO measurement integrated into trial
- PRO Instruments: SF-12; MOS Vitality scale; modified BCPT sx; 10-item CES-D; utility rating scale; MOS sexual functioning
- Key PRO Questions: Expect no difference in physical and mental health, but differential patterns of symptoms

NSABP – 35: Design Specifics

- Accrual: 3000 patients/ first 1175 on PRO study
- Postmenopausal, stratified by age 60 or less
- DCIS without invasion
- ER or PR positive/IHC
- Lumpectomy/Negative margins
- Radiation therapy
- Start treatment within 84 days
- Double masked, placebo controlled, treatment for 5 years
- PRO assessments prior to treatment and q6 months

NSABP B-35

- Primary Aim
  - Compare anastrozole to tamoxifen in preventing the occurrence of breast cancer in postmenopausal women following lumpectomy and radiation therapy for DCIS

NSABP B-35

- Secondary Aims
  - Invasive breast cancer
  - Ipsilateral cancer recurrence
  - Contralateral breast cancer
  - QOL
  - Osteoportic fractures
  - DFS
  - OS

NSABP B-35: PRO Hypotheses

- Primary
  - No difference in MCS or PCS of SF-12
  - Hot flashes > with tamoxifen and most pronounced in <60 years
- Secondary
  - Vaginal dryness and sexual functioning worse with anastrozole
  - Better quality adjusted survival with anastrozole

NSABP B-35: Comments

- PRO instruments build on prior NSABP prevention and treatment trials, especially P-1 and P-2
- Shortened measures and specific hypotheses derive from prior work with these scales
- Drug toxicities in this patient population and PRO effects well-understood

NSABP B-35: Other features

- Integrated into trial design with specific, relevant questions
- Sample size targeted to the PRO question
- Compliance monitoring prospectively, with identification of problem sites
- PRO data collection included in institutional performance evaluation

NSABP B-35 Compliance Report as of 8/09/06

B-35 QOL Compliance	# Patients with Form Expected	# Patients with QOL/QMD form Submitted	Percentage of patients with QOL/QMD submitted	Percentage of Patients with QOL Submitted	Percentage of submitted forms that were QMD
Baseline	1275	1275	100	100	0
6 Months	1257	1252	100	94	5
12 Months	1243	1223	98	92	6
18 Months	1225	1125	92	87	6
24 Months	892	727	82	76	6
30 Months	435	314	72	67	7
36 Months	134	79	59	55	6
42 Months	0	0			
48 Months	0	0			
54 Months	0	0			
60 Months	0	0			
66 Months	0	0			
72 Months	0	0			

NSABP B-35: Conclusions

- PRO questions important with excellent survival and poor tolerance of side effects in DCIS—a prevention setting!
- Differences in treatment outcomes are likely small; PRO outcomes important for ultimate treatment decisions
- Symptom patterns different across the two agents
- PRO assessment strengthened by double blind, placebo controlled design

What have we found out?

- Design
  - Phase III studies with QOL as a primary or secondary endpoint may be successful
  - Equivalence studies are possible though require large sample sizes
  - “Companion studies” can be successful with a large enough N - however, they are not optimal

What have we found out?

- Assessment
  - Standardized questionnaires are available and feasible in the cooperative group, Phase III setting
  - Availability of norms, comparable data, and well-supported clinical meaningfulness guidelines eases interpretation
  - Multiple measures and multiple times points can cloud interpretation; special concern for missing data and staff burden

What have we learned?

- Analysis
  - Many analytic approaches for addressing missing data are available
    - “Imputing” data points
    - Only include patients with complete data
    - Subgroups depending on data completion patterns
    - Statistical analyses:
      - Data missing at random?
      - Mixed linear model, pattern mixture model

What have we learned?

- Minimizing missing data is the simplest approach
  - Match patient characteristics and times of assessment
  - Many measurement points and long-term assessments of severely ill patients will likely result in considerable missing data
  - Include quality control systems to monitor PRO data submission

What have we learned?

- Publication
  - Biological explanations and clearcut implications for clinical practice may increase a study's appeal
  - Breaking a therapeutic study into several papers is probably necessary; a series of papers in the same journal examining, e.g., outcomes re: treatment, QOL, economics, etc. would be optimal

Optimal Conditions for Inclusion of PROs in Phase III Trials

- Integrated planning/inclusion of PRO endpoints into trial protocol
- PRO endpoints selected for ‘value-added’ for clinicians and patients—will the PRO data make a difference at the end of the trial?
- A priori hypotheses are critical

Optimal Strategies for Obtaining High Quality Data

- Delinquency in PRO data are treated like any other data in monitoring clinical site performance
- Regular review of data submission and compliance during the trial
- Monitoring of sites with counseling of those with delinquency during the trial
- Information support for collection of PRO data, FAQs, as well as modest financial incentives

Optimal Analysis and Reporting

- PRO endpoints are analyzed and included as either a secondary or primary endpoint of the trial when first reported
- More detailed elaboration of PRO endpoints in a companion manuscript within the same journal

What do we conclude?

- There has been significant progress in the successful inclusion of PRO endpoints in phase III clinical trials
- Each particular trial has its challenges
- Over time, increasing resources have been committed to this activity and the quality of results reflect varying commitment of the parent clinical trial group