

*Journal Club*

Phase II Studies:  
a brief example from the  
literature and overview of design

Elizabeth Garrett-Mayer, PhD  
and

Elizabeth A. Griffiths, MD

# Outline

- Clinical trial phase review
- Brief overview of Prostate Cancer
  - Statistics/diagnosis
  - Standard of care
  - Ongoing research
- Randomized Phase II trials- details of design and agenda
- A good example of a randomized phase II trial:
  - Oudard et al : Study design, details and results
- Discussion

# Phases of Study Treatment: Definitions

- Phase I trials:
  - identify the dose range that is well tolerated
  - helps to identify common or marked side effects
  - Usually involves a very small number of patients
  - Usually does NOT have a control group.
- Phase II trials:
  - provide preliminary information on whether a treatment is efficacious
  - provide preliminary data about the relationship between dose and efficacy.
  - **Often controlled but have too few patients to detect any but the largest differences in treatment effect.**
  - Often unblinded.
- Phase III trials:
  - provide definitive evidence of efficacy
  - detail common side effects
  - many participants.
  - often blinded.

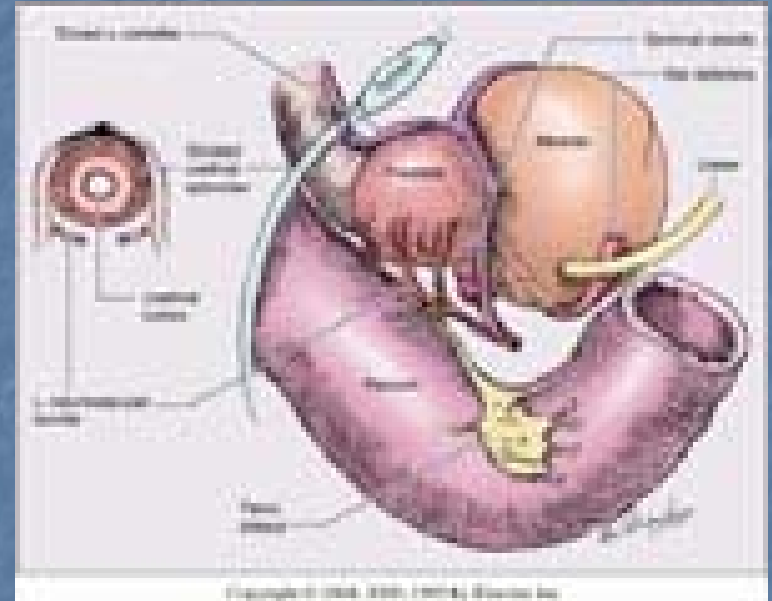
*From: Fletcher RH, SW Fletcher and EH Wagner. Clinical Epidemiology: the essentials. Baltimore: Williams and Wilkins, 1996.*

# Variations on a theme

- Phase I and II trials can be combined as a phase I/II to address both questions simultaneously. This is particularly useful with completely novel compounds where toxicity and efficacy may be assessed together.
- Phase II and phase III trials can be combined (more later....)
- Phase IV trials or post marketing assessments are designed to assess more subtle side effects and toxicities beyond the confines of a tightly controlled clinical trial.

# Prostate Cancer: An Overview

- The most commonly diagnosed life threatening malignancy in men (220,900 cases and 28,900 deaths per year).
- 29% of men between 30 and 40 and 64% of men between 60 and 70 will have small prostate cancers
- The lifetime risk of prostate cancer is 1 in 6 and the risk of death 1 in 30
- Risk factors include age, family history (increases risk to  $\frac{1}{4}$  with affected father or brother), diet, lifestyle and ethnicity



# Prostate Cancer

- Strategies for prevention are not well delineated.
- Screening for this disease includes annual PSA and digital rectal exam and seems to detect smaller cancers localized to the prostate
- Diagnosis is via transrectal ultrasound guided core biopsy
- Prognosis is determined by stage, Gleason score and PSA at diagnosis.

# Standard of Care

## ■ Local disease

- Observation
- radical prostatectomy (RP)
- external beam radiation(xrt) and brachytherapy.  
(addition of androgen suppression to xrt may improve survival.)

## ■ Advanced disease

- 1<sup>st</sup> line- androgen suppression (either chemical or physical castration). *All prostate cancers will ultimately become androgen independent and progress.*
- 2<sup>nd</sup> line- docetaxel based chemotherapy or a clinical trial

# Hormone Refractory Prostate Cancer

- The second leading cause of cancer death in men in the west.
- Median Survival is 1yr
- Systemic chemotherapy appears to improve quality of life and therapy with Docetaxel based regimens has recently been demonstrated to prolong survival from 15 to 18 months.
- Given the relatively poor response of prostate cancer to current chemotherapeutic regimens numerous clinical trials are under way.

# Mechanisms of Hormone Refractoriness

- Hormone refractory prostate cancers appear to over-express the Androgen Receptor (AR) relative to hormone sensitive tumors. In theory this alteration might allow the tumor to respond to much lower levels of endogenous androgens.
- Activating mutations in the AR gene may alter ligand specificity such that non-androgenic steroids as well as antiandrogens may activate the mutant AR.
- Other mutations result in changes in receptor affinity
- Downstream molecular changes in regulatory genes may prevent prostate cancer cells from responding to apoptotic signals.
- Other mechanisms have also been described

# Current standard of care in 1<sup>st</sup> line metastatic hormone refractory prostate cancer:

- Docetaxel 75mg/m<sup>2</sup> given every 3 weeks with prednisone 5mg po bid. Median survival 16.5 vs 18.9mo (p<0.0005), PSA RR (PRR) 45%<sup>1</sup>
- Docetaxel 60mg/m<sup>2</sup> D2 with Estramustine 280 po tid D1-5. Median survival 18 vs 15mo (p<0.008), PRR 50%<sup>2</sup>
- Docetaxel 36mg/m<sup>2</sup> with calcitriol 0.5mcg/kg given weekly for 6/8 wks. PRR 81%<sup>3</sup>
- Mitoxantrone 12mg/m<sup>2</sup> every 21 days with prednisone 5mg po bid. Improves quality of life but NOT overall survival or disease free survival in two randomized controlled trials. <sup>4</sup>

1. Eisenberger M et al. NEJM Oct 2004; 351:1502-12.

2. Petrylak D. P et al. NEJM Oct 2004; 351:1513-20.

3. Tannock IF et al. JCO 1996; 14(6):1756-64.

4. Beer TM et al. semin oncol 2001; 28(4):49-55.

# Open Trials at Hopkins

- [ECALGB90401](#): Randomized, Double-Blinded Placebo Controlled Phase III Trial Comparing Docetaxel and Prednisone with and w/o Bevacizumab in Men with Hormone Refractory Prostate Cancer
- [J0458](#): A Phase III Randomized, Open-Label Study of prostate cancer vaccine Versus Docetaxel and Predinisonone in Patients with Metastatic Hormone-Refractory Prostate Cancer who are Chemotherapy-Naive
- [J0310](#): Phase II Study of Infliximab (remicade) in Androgen-Independent Prostate Cancer with Bone Metastases
- [J0513](#): Phase II Trial of 17-AAG in Patients with Hormone Refractory Metastatic Prostate Cancer
- [J0467](#): Phase I Trial with a Combination of Docetaxel + 153 Sm-EDTMP (Samarium 153) in Patients with Hormone-Refractory Prostate Cancer

# Randomized Phase II Trials

- As seen from the prior slides, the current standard of care for hormone refractory prostate cancer has been delineated by several recent (2004) randomized phase III trials.
- The trial we review today is a randomized phase II trial which **does not** present new information as to the strategy for treating hormone refractory prostate cancer.
- **It does however provide a context through which to discuss the design, focus and expected results of a good phase II trial.**

# Common quandary these days

- Targeted therapies combined with standard, already used chemotherapies
- For many molecularly targeted agents, there is little evidence of single agent activity, but thought that combination of agents with chemo will be effective.
- Phase I trial shows tolerability.
- What next? Phase III?
- Phase III is bad idea in most cases: need preliminary clinical efficacy evidence.
- Recent leaps from Phase I to Phase III have proven costly (and haven't panned out)\*.

\*Moore et al. Gemcitabine vs. Matrix Metalloproteinase inhibitor (JCO, 2003)  
Van Cutsem et al. Gemcitabine + tipifarnib vs. Gemcitabine + placebo (JCO, 2004)

# Why randomized phase II?

## ■ Classic phase II studies:

- Single arm study where results are compared to 'historical control' rate.
- Problem: this is not always 'satisfying'
  - Requires patient populations to be comparable
  - Might not have information to derive control rate (e.g. disease progression is of interest and not response)

## ■ Comparative randomized studies (phase III):

- Allow us to compare two arms
- Problem:
  - Large sample size (more than twice a single arm study)
  - Costly
  - Large undertaking based on scant preliminary data

# Why randomized phase II?

- Want to explore efficacy
- Not willing to invest in phase III (yet)
- Want some “control” or “prioritization”
- Primarily two different kinds of randomized phase II studies
  - Phase II selection design (prioritization)
  - Phase II designs with reference control arm (control)
- Also phase II/III studies

# Common design of randomized phase II study

- Two parallel one arm studies (classic case)
- **Do not directly compare arms to each other.**
- Compare each to “null rate”
- Example: null response = 0.25, alternative response=0.50, alpha=0.10, power=0.90.
  - Two parallel one-arm studies:
    - Test each treatment to see if it is better than null rate
    - For two arm study, need N=70 patients (35 per arm)
  - Comparative study:
    - Test to see if one treatment is better than the other treatment
    - For two arm study, need N=170 patients (85 per arm)

# Classic Randomized Phase II designs

- Phase II selection designs (Simon, 1985)
  - “pick the winner”
  - 90% chance of choosing better arm so long as difference in response rates is  $>15\%$ .
  - Appropriate to use when:
    - Selecting among NEW agents
    - Selecting among different schedules or doses
  - NOT appropriate when
    - Trying to directly compare treatment efficacies (not powered)
  - Uses 2+ Simon two-stage designs
    - Each arm is compared to a null rate
    - Must satisfy efficacy criteria of Simon design
    - Move the “winner” to phase III
    - Only have to pick winner if more than one arm shows efficacy
  - Can be used when the goal is prioritizing which (if any) experimental regimen should move to phase III when no a priori information to favor one.

# Classic Randomized Phase II designs

- Randomized Phase II designs with reference arm
  - Includes reference arm to ensure that historical rate is “on target”
  - Reference arm is not directly compared to experimental arm(s) (due to small N)
  - Can see if failure (or success) is due to incomparability of patient populations
  - Problem: if it turns out that historical control rate used is very different from what is observed in reference arm, then trial should be repeated\*

\*Herson and Carter, Statistics in Medicine (1986)

# Phase II/III studies

- Several versions {Schaid (1988), Storer (1990), Ellenberg and Eisenberger (1985), Schaid and Heller (2002)}
- General idea
  - Begin with randomized phase II study
  - Randomize to control arm & experimental arm(s)
  - If some threshold of efficacy is met, continue to phase III sample size for direct comparison
- Benefits:
  - Allow use of phase II data in phase III inference
  - Minimize delay in starting up phase III study
  - Uses concurrent control
- Cons:
  - The sample size for the phase II part is approximately twice as large as would be needed for standard phase II
  - Need phase III infrastructure developed even if it stops early.
- **Would be useful if MOST phase II studies showed efficacy**
- Really, these could be considered phase III designs with very aggressive early stopping rules.

# Other randomized Phase II designs?

Lots of randomized studies are calling themselves randomized phase II studies these days:

- If outcome of interest is surrogate
  - Correlative (biomarker)
  - Clinical (response)
- If sample size is relatively small but direct comparison is made (current study?)
- If study is comparative, but is not definitive for whatever reason.

## Multicenter Randomized Phase II Study of Two Schedules of Docetaxel, Estramustine, and Prednisone Versus Mitoxantrone Plus Prednisone in Patients With Metastatic Hormone-Refractory Prostate Cancer

*Stéphane Oudard, Eugeniu Banu, Philippe Beuzevot, Eric Voog, Louis Marie Dourthe, Anne Claire Hardy-Bessard, Claude Linassier, Florian Scotté, Adela Banu, Yvan Coscas, François Guinet, Marie-France Poupon, and Jean-Marie Andrieu*

- 130 metastatic HRPC patients randomized to one of three arms:
  - Docetaxel in one of two regimens (Arm A = 70mg/m<sup>2</sup> on d 2 or Arm B = 35mg/m<sup>2</sup> on d 2,9 every 21 d) with estramustine (280mg po tid d 1-5, 8-12) and prednisone (10mg po daily) or
  - Arm C = mitoxantrone (12mg/m<sup>2</sup> every 21d) and prednisone (10mg po daily)
- 127 patients were evaluated.
- Outcomes assessed were PSA response (primary end point) and safety

# Randomized Phase II design

- 42 patients per arm
- Null rate of response=30%, alternative rate=60%
- What are they powered to see?
- With 80% power and alpha of 0.05, they can DIRECTLY compare each arm to the other.
- Why randomized phase II?
- Unclear. Maybe because....
  - Looking for huge difference (30% vs. 60%)
  - They don't consider response rate definitive (reasonable: survival would be a better outcome for phase III study).
- Also: they state that "Simon design" was used. But, this seems inconsistent with their stated design.

**Table 1.** Baseline Characteristics of Patients According to Random Assignment (N = 127)

Characteristic	Arm A* (n = 43)		Arm B† (n = 42)		Arm C‡ (n = 42)	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
Age, years						
Median	68		68		70	
Interquartile range	52-91		51-79		52-85	
ECOG performance status						
0	17	40	25	59	20	48
1	19	44	13	31	11	26
2	7	16	4	10	11	26
Gleason score						
2-4	2	5	0	0	1	2
5-6	10	23	5	12	10	24
7-10	30	70	37	88	28	67
Unknown	1	2	0	0	3	7
Time from diagnosis of prostate cancer to random assignment, months						
Median	33		33		47	
Interquartile range	3-219		5-151		6-150	
Time from start of hormonal treatment to random assignment, months						
Median	16		27		25	
Interquartile range	2-116		2-89		1-118	
Tumor-related symptoms						
Without bone pain	12	28	12	29	11	26
Bone pain	28	65	24	57	30	72
Unknown	3	7	6	14	1	2
Analgesic use at entry						
Analgesic treatment	24	56	21	50	25	60
No analgesic treatment	16	37	14	33	16	38
Unknown	3	7	7	17	1	2
Serum PSA, ng/mL						
Median	71		69.5		77.7	
Interquartile range	1.9-2818		0.01-2416		0.41-1840	
Hemoglobin level, mg/dL						
Median	12.9		12.4		12.9	
Interquartile range	9.9-17		8.2-15.7		8.7-15.6	
Number of organs involved						
One	27	63	28	67	27	64
Two	16	37	11	26	14	33
≥ Three	0	0	3	7	1	3
Sites of metastases						
Bone	38	88	39	93	41	98
Lymph nodes	16	37	11	26	13	31
Other sites	5	12	8	19	3	7
Type of previous hormonal regimen						
Total androgen blockade	36	84	36	86	35	83
Number of previous hormonal regimens						
One	30	70	26	62	32	76
Two	11	25	12	29	10	24
Three	2	5	4	9	0	0
Other previous anticancer therapy						
Surgery	6	14	10	24	8	19
Radiotherapy	10	23	9	21	7	17

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PSA, prostate-specific antigen.

\*Docetaxel-estramustine-prednisone on day 2 every 21 days.

†Docetaxel-estramustine-prednisone on days 2 and 9 every 21 days.

#Mitoxantrone-prednisone every 21 days.

**Table 2.** PSA Response Rates

PSA Response	Arm A* (n = 43)		Arm B† (n = 42)		Arm C‡ (n = 42)	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
≥ 50% §	29	67	26	63	7	18
≥ 75% §	22	51	16	39	3	8
Normalisation, < 4 ng/mL	10	23	7	17	1	2

NOTE. One patient in arm B and three patients in arm C could not be evaluated for PSA response because of a baseline PSA level < 4 ng/mL.

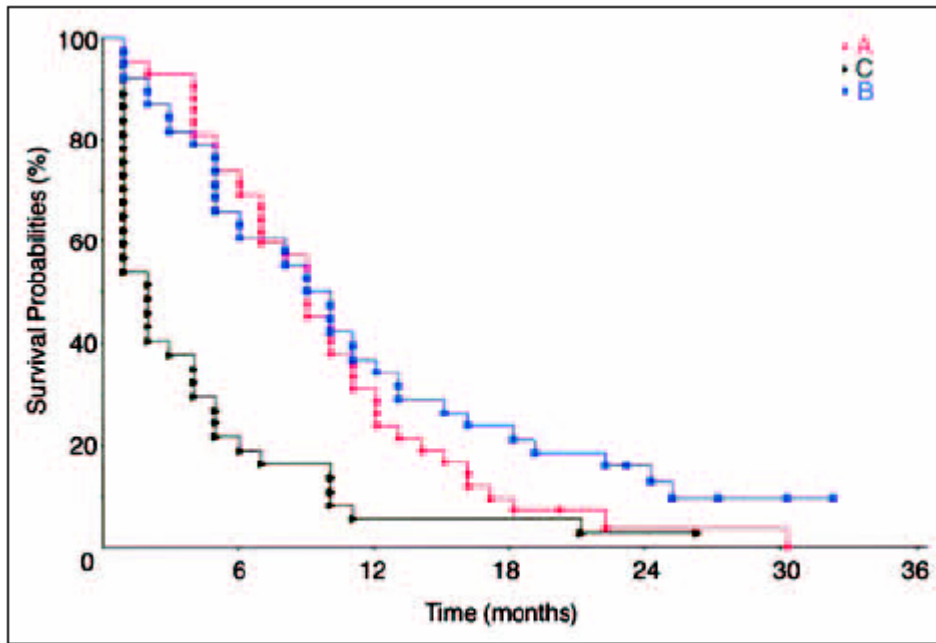
Abbreviation: PSA, prostate-specific antigen.

\*Docetaxel-estramustine-prednisone on day 2 every 21 days.

†Docetaxel-estramustine-prednisone on days 2 and 9 every 21 days.

‡Mitoxantrone-prednisone every 21 days.

§ $P < .0001$ .

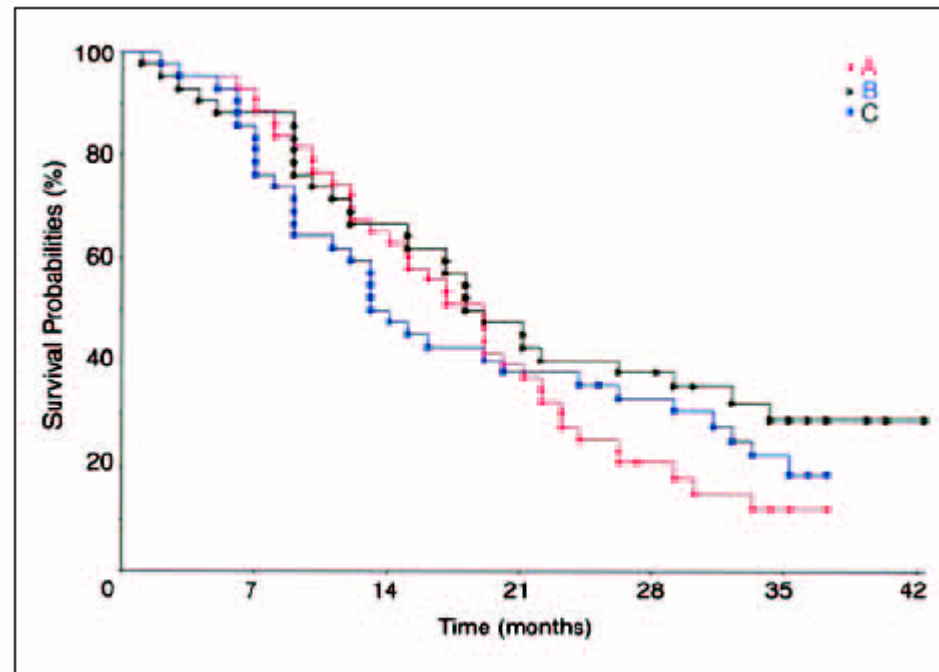


**Fig 1.** Kaplan-Meier plot of time to prostate-specific antigen progression for each treatment arm.

Figure 1:  
 Median time to progression was more than five times longer with the DEP arms compared to the MP arm 8.8months (CI 6.9-10.8) - 9.3 months (CI 7.5-11.1) vs 1.7 months (CI 0.7 to 2.7) (p= 0.000001) the DEP groups also had a longer duration of response (8 and 8.3 vs 6.4 months)

Figure 2:

This study was not powered to demonstrate a survival difference and at the time of analysis there was NO statistically significant survival difference between the two groups.



**Fig 2.** Kaplan-Meier plot of overall survival in each treatment arm.

**Table 3.** Clinical Benefit Response

Characteristic	Arm A* (n = 43)		Arm B† (n = 42)		Arm C‡ (n = 42)	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
Pain control	10	23	9	21	7	17
Analgesic consumption	15	35	10	24	6	14
Improved pain index,§ 1 + 2	17	40	12	29	7	17
Improved ECOG PS	26	60	20	48	12	28
Improved clinical benefit,§ 3 + 4	14	33	10	24	9	21

Abbreviation: ECOG PS, Eastern Cooperative Group performance status.

\*Docetaxel-estramustine-prednisone on day 2 every 21 days.

†Docetaxel-estramustine-prednisone on days 2 and 9 every 21 days.

‡Mitoxantrone-prednisone every 21 days.

§Evaluated according to the validated modified McGill pain questionnaire.

|| $P = .01$ .

**Table 4.** Toxicity: Severe Adverse Events (grade 3 or 4)

Toxicity	Arm A* (n = 43)		Arm B† (n = 42)		Arm C‡ (n = 42)	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
Granulocytopenia	16	37	0	0	20	48
Granulocytopenic fever	0	0	0	0	3	7
Anemia	1	2	0	0	3	7
Thrombocytopenia	0	0	1	2	1	2
Nausea	1	2	0	0	0	0
Vomiting	1	2	0	0	0	0
Diarrhea	3	7	0	0	0	0
Thrombosis	3	7	3	7	0	0

\*Docetaxel-estramustine-prednisone on day 2 every 21 days.

†Docetaxel-estramustine-prednisone on days 2 and 9 every 21 days.

‡Mitoxantrone-prednisone every 21 days.

**Table 5.** Relative Event Rate

Parameter	A v B		A v C		B v C		<i>P</i>
	RER	95% CI	RER	95% CI	RER	95% CI	
Time to PSA progression	1.26	0.85 to 1.89	0.44	0.25 to 0.76	0.35	0.20 to 0.60	.00001
Overall survival	1.43	0.89 to 2.31	1.08	0.66 to 1.76	0.75	0.46 to 1.21	.13
Time on primary treatment	1.26	0.78 to 2.04	0.49	0.25 to 0.97	0.39	0.20 to 0.76	.0005

Abbreviations: RER, relative event rate; PSA, prostate-specific antigen.

**Table 6. Association of Baseline Factors With Overall Survival  
Univariate and Multivariate Analysis**

Factor	<i>P</i>
<b>Univariate</b>	
ECOG performance status	.0001
Hemoglobin	.0001
Prior duration of hormone therapy	.002
Bone pain at presentation	.0001
<b>Multivariate</b>	
ECOG performance status	.0001
Hemoglobin	.006

Abbreviation: ECOG, Eastern Cooperative Oncology Group.

# Discussion 1

- They got a “home-run:” with small sample sizes, they could conclude significant differences in response
- Their post-hoc multiple comparison adjustment is sketchy: should have been done as part of study design.
- Still: had this been “the” study of these agents, this would have been strong evidence to take arm A or arm B forward to Phase III
- But poor design! Which would we take forward?!

# Discussion 2

- Randomized phase II studies should not be considered “short-cut” way to comparative study
  - “i want to do a comparative study, but there is no way i can get 200 patients....”
- Should be used for one of two goals:
  - 1) Prioritization of regimens for phase III evaluation
  - 2) Validation of control rate
- Current research climate: many candidates! Critical to screen these because we cannot take so many forward to Phase III.
- **Important:** these studies need to protect the ability to perform definitive phase III trials

# Discussion 3

- This trial presents findings which are better detailed in other trials.
- It essentially confirms in an independent patient population similar results.
- The authors were lucky: with a small study that was designed to look for a big difference, they actually saw the big difference
- In general, need very strong evidence to make sweeping conclusions from randomized phase II studies.