

ESMO CONFERENCE  
11TH WORLD CONGRESS ON  
**Gastrointestinal  
CANCER**  
24-27 JUNE 2009  
BARCELONA, SPAIN

Imedex

## Issues and challenges in choosing endpoints for phase III clinical trials

Marc Buyse, ScD  
IDDI, Louvain-la-Neuve, and  
University of Hasselt, Diepenbeek, Belgium  
[marc.buyse@iddi.com](mailto:marc.buyse@iddi.com)

## Outline

- Choosing a primary endpoint
- Pros and cons of different endpoints
- Why survival should no longer be a gold standard in cancer trials
- Validating surrogate endpoints for survival

## Endpoints

All trials have a « primary » endpoint, which is used to

- calculate the number of patients required
- drive the interpretation of the trial results

Most trials include additional endpoints

- secondary endpoints
- exploratory (including « translational ») endpoints

## Why are endpoints important?

Example in advanced cancer:  
Randomized trial for patients receiving cytotoxic chemotherapy ± G-CSF, a growth factor that accelerates neutrophil recovery

Patients receiving cytotoxic therapy

### Choice of endpoint, expected treatment effect, and number of patients required

PARAMETER	VALUES		RELATIVE RISK	NR OF PTS REQUIRED
	NO G	G		
Mean duration of neutropenia	9 days	3 days	0.33	20
% with febrile neutropenia	60%	30%	0.50	100
% with treatment delay	40%	20%	0.50	200
% with documented infection	15%	7.5%	0.50	500
% dying from infection	5%	2.5%	0.50	2000
Response rate			0.75 ?	Hundreds
Time to tumor progression			0.85 ?	Thousands
Overall survival			0.95 ?	Thousands

### Choice of endpoint, expected treatment effect, and number of patients required

PARAMETER	VALUES		RELATIVE RISK	NR OF PTS REQUIRED
	NO G	G		
Mean duration of neutropenia	9 days	3 days	0.33	20
% with febrile neutropenia	60%	30%	0.50	100
% with treatment delay	40%	20%	0.50	200
% with documented infection	15%	7.5%	0.50	500
% dying from infection	5%	2.5%	0.50	2000
Response rate			0.75 ?	Hundreds
Time to tumor progression			0.85 ?	Thousands
Overall survival			0.95 ?	Thousands

Choice of endpoint, expected treatment effect, and number of patients required

PARAMETER	Increasing clinical relevance	Decreasing treatment benefit	RELATIVE RISK	NR OF PTS REQUIRED
Mean time to progression	0.33	0.33	0.33	20
% with grade 3-4 neutropenia	0.50	0.50	0.50	100
% with grade 3-4 diarrhea	0.50	0.50	0.50	200
% with grade 3-4 nausea	0.50	0.50	0.50	500
% with grade 3-4 vomiting	0.50	0.50	0.50	2000
Response rate	0.75 ?	0.75 ?	0.75 ?	Hundreds
Time to tumor progression	0.85 ?	0.85 ?	0.85 ?	Thousands
Overall survival	0.95 ?	0.95 ?	0.95 ?	Thousands

Choice of endpoint, expected treatment effect, and number of patients required

PARAMETER	Increasing clinical relevance	Decreasing treatment benefit	RELATIVE RISK	NR OF PTS REQUIRED
Mean time to progression	0.33	0.33	0.33	20
% with grade 3-4 neutropenia	0.50	0.50	0.50	100
% with grade 3-4 diarrhea	0.50	0.50	0.50	200
% with grade 3-4 nausea	0.50	0.50	0.50	500
% with grade 3-4 vomiting	0.50	0.50	0.50	2000
Response rate	0.75 ?	0.75 ?	0.75 ?	Hundreds
Time to tumor progression	0.85 ?	0.85 ?	0.85 ?	Thousands
Overall survival	0.95 ?	0.95 ?	0.95 ?	Thousands

Overall survival is the most important endpoint in advanced cancer...

... but that does not make it a good primary endpoint for clinical trials!

Median survival of patients with metastatic cancer

Up to 1990 (5FU): < 12 months

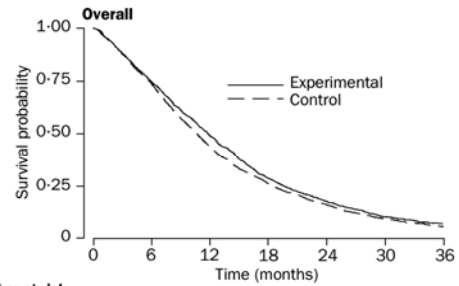
Circa 1995 (optimization of 5FU): ~ 12-15 months

Circa 2000 (oxaliplatin, irinotecan): ~ 15-18 months

Today (bevacizumab, cetuximab, panitumumab): > 18 months\*

\* longer in targeted subsets (e.g. wild type Kras)

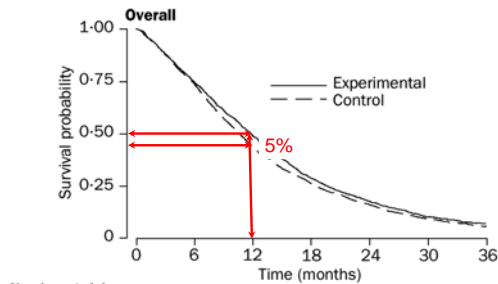
Survival benefits of 5FU optimization



Number at risk		2031	1504	956	524	287	147	89
Experimental		2031	1504	956	524	287	147	89
Control		1760	1284	739	408	229	112	53

Ref: Buyse et al, Lancet 2000;356:373.

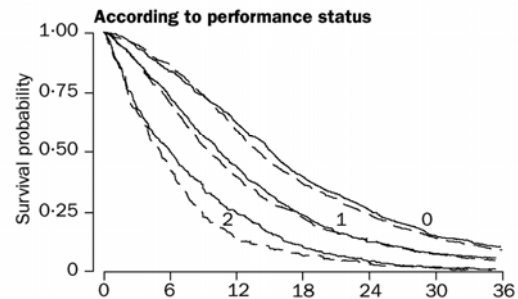
Survival benefits of 5FU optimization



Number at risk		2031	1504	956	524	287	147	89
Experimental		2031	1504	956	524	287	147	89
Control		1760	1284	739	408	229	112	53

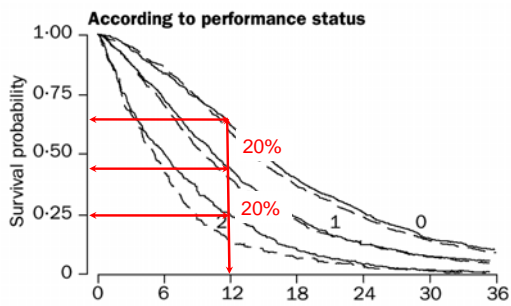
Ref: Buyse et al, Lancet 2000;356:373.

Survival benefits of 5FU optimization



Ref: Buyse et al, Lancet 2000;356:373.

### Survival benefits of 5FU optimization



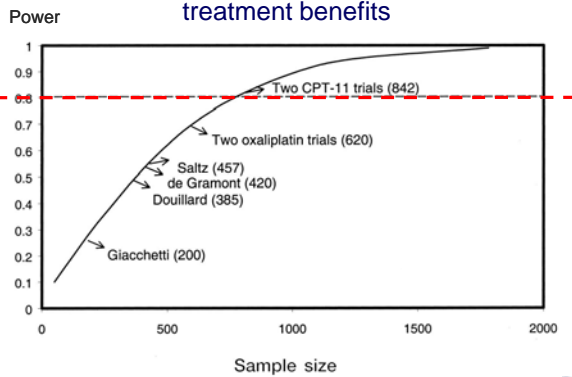
Ref: Buyse et al, Lancet 2000;356:373.

### Overall survival is insensitive to real treatment benefits

- **Metastatic colorectal cancer:** 4 phase III clinical trials used for registration of oxaliplatin or irinotecan added to standard first-line therapy with 5FU+leucovorin
- Power calculated for a 25% hazard reduction (i.e. a hazard ratio of 0.75) at a two-sided significance level ( $\alpha$ ) of 0.05.

Ref: DiLeo, Bleiberg and Buyse, JCO 2005;23:8664.

### Overall survival is insensitive to real treatment benefits



### Reasons for better sensitivity of PFS as compared with OS

- Larger number of events at same follow-up time
- Larger treatment effect on PFS than on OS
- PFS less affected by competing risks (especially in elderly populations)
- PFS unaffected by effective rescue therapies and successive treatment lines

### Tumor response

Pros	Cons
1. Measured early (3 – 6 months)	1. Responses infrequent
2. Measured easily	2. Insensitive to disease stabilizations (cytostatics)
3. Assessment can be reviewed blindly by expert committee	3. Assessment prone to error and/or bias if not reviewed
4. Reflects biological activity	4. Disease not always measurable
5. Large treatment effects	5. Limited impact on survival

### Progression-free survival

Pros	Cons
1. Reflects control of disease process	1. Assessment subjective
2. Unaffected by competing risks of death	2. Assessment potentially biased to allow for change in therapy
3. Very sensitive to differences in treatment efficacy	3. Assessment can be reviewed only after changes in therapy
4. Possible impact on survival	
5. Closely related to quality of life	

### Survival

Pros	Cons
1. Most meaningful	1. Hard to affect, therefore large sample sizes needed
2. Most objective	2. Measured late
	3. Affected by second-line treatments
	4. Affected by competing risks
	5. Insensitive to short-term benefits

### Choice of endpoint, expected treatment effect, and number of patients required

PARAMETER	VALUES		RELATIVE RISK	NR OF PTS REQUIRED
	OLD	NEW		
Tumor response	20%	40%	OR = 0.38	70
Progression-Free Survival	9 mos	12 mos	HR = 0.75	700
Overall survival	18 mos	21 mos	HR = 0.85	2100

- ### Requirements for ideal endpoint
- Ideal endpoint should
- have clear definition
  - be easy to measure
  - have little opportunity for ascertainment bias
  - capture all clinically relevant events
  - be sensitive to real treatment benefits
  - be observed as early and in as many patients as possible

- ### Requirements for ideal endpoint
- | Ideal endpoint should                                      | OS | PFS |
|--|----|-----|
| • have clear definition                                    | +  | -   |
| • be easy to measure                                       | +  | -   |
| • have little opportunity for ascertainment bias           | +  | -   |
| • capture all clinically relevant events                   | -  | +   |
| • be sensitive to real treatment benefits                  | -  | +   |
| • be observed as early and in as many patients as possible | -  | +   |

- ### Alleged problems with PFS
- Measurement errors  
*Do they matter?*
  - Frequency of measurement  
*Does it matter?*
  - Patients leaving trial before progression  
*Should these observations be censored?*
  - Assessment bias (in non double blind trials)  
*Is central review needed?*
- Ref: Sargent and Hayes, JCO 2008;26:1922.

### Is central review needed?

Trial	Sample Size	Per Central Review		Per Local Review	
		HR	95% CI	HR	95% CI
<b>Renal cell carcinoma</b>					
Sorafenib v placebo <sup>17</sup>	903	0.44	0.35 to 0.55	0.51	0.43 to 0.60
Sunitinib v interferon alpha <sup>18</sup>	750	0.42	0.32 to 0.54	0.42	0.33 to 0.52
<b>Colorectal cancer</b>					
Panitumumab plus best supportive care v best supportive care <sup>19</sup>	463	0.54	0.44 to 0.66	0.39	0.32 to 0.48
<b>Breast cancer</b>					
Lapatinib plus capecitabine v capecitabine <sup>20</sup>	324	0.49	0.34 to 0.71	0.59	0.42 to 0.84
Bevacizumab plus capecitabine v capecitabine <sup>21</sup>					
Bevacizumab plus paclitaxel v paclitaxel <sup>16, 11</sup>	722	0.42	0.34 to 0.52	0.48	0.39 to 0.61
Trastuzumab plus capecitabine v capecitabine <sup>22</sup>	752	NR		NR	
Median PFS		5.8 v 4.2		5.3 v 3.8	

Ref: Dodd et al, JCO 2008;26:3791.

## Is central review needed?

Trial	Sample Size	Per Central Review		Per Local Review	
		HR	95% CI	HR	95% CI
Renal cell carcinoma					
Sorafenib v placebo <sup>17</sup>	903	0.44	0.35 to 0.55	0.51	0.43 to 0.60
Sunitinib v interferon alpha <sup>18</sup>	750	0.42	0.32 to 0.54	0.42	0.33 to 0.52
Colorectal cancer					
Panitumumab plus best supportive care v best supportive care <sup>19</sup>	463	0.51	0.44 to 0.66	0.39	0.32 to 0.48
Breast cancer					
Lapatinib plus capecitabine v capecitabine <sup>20</sup>	324	0.49	0.34 to 0.71	0.59	0.42 to 0.84
Bevacizumab plus capecitabine v capecitabine <sup>21</sup>	462	0.98	0.77 to 1.25	NR, value not significantly different from 1	
Bevacizumab plus paclitaxel v paclitaxel <sup>10,11</sup>	722	0.42	0.34 to 0.52	0.48	0.39 to 0.61
Ixabepone plus capecitabine v capecitabine <sup>22</sup>	752	NR		NR	
Median PFS		5.8 v 4.2		5.3 v 3.8	

Ref: Dodd et al, JCO 2008;26:3791.

## Is central review needed?

Trial	Sample Size	Per Central Review		Per Local Review	
		HR	95% CI	HR	95% CI
Renal cell carcinoma					
Sorafenib v placebo <sup>17</sup>	903	0.44	0.35 to 0.55	0.51	0.43 to 0.60
Sunitinib v interferon alpha <sup>18</sup>	750	0.42	0.32 to 0.54	0.42	0.33 to 0.52
Colorectal cancer					
Panitumumab plus best supportive care v best supportive care <sup>19</sup>	463	0.51	0.44 to 0.66	0.39	0.32 to 0.48
Breast cancer					
Lapatinib plus capecitabine v capecitabine <sup>20</sup>	324	0.49	0.34 to 0.71	0.59	0.42 to 0.84
Bevacizumab plus capecitabine v capecitabine <sup>21</sup>	462	0.98	0.77 to 1.25	NR, value not significantly different from 1	
Bevacizumab plus paclitaxel v paclitaxel <sup>10,11</sup>	722	0.42	0.34 to 0.52	0.48	0.39 to 0.61
Ixabepone plus capecitabine v capecitabine <sup>22</sup>	752	NR		NR	
Median PFS		5.8 v 4.2		5.3 v 3.8	

Ref: Dodd et al, JCO 2008;26:3791.

## Are there good surrogates for OS?

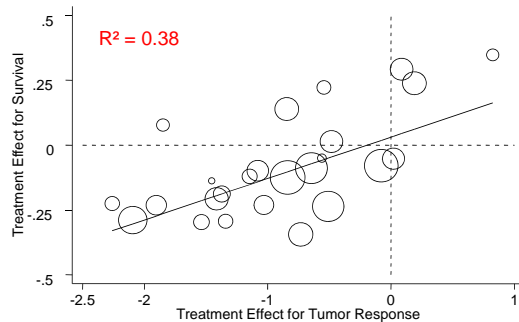
- Tumor response is not a good surrogate for survival in solid tumors (even though response is predictive of survival)
- DFS is an excellent surrogate for survival in early colorectal and gastric cancer
- PFS is a good surrogate for OS in some solid tumors (eg colorectal, lung and ovarian cancer) but not in others (eg breast and prostate cancer)

## 4 meta-analyses of experimental 5-FU

	Response rate	Median survival
Control	220/1814=12%	12 months
Experimental	479/2084=23%	13 months
	P<0.000000001	P=0.003

Ref: Buyse et al, Lancet 2000;356:373.

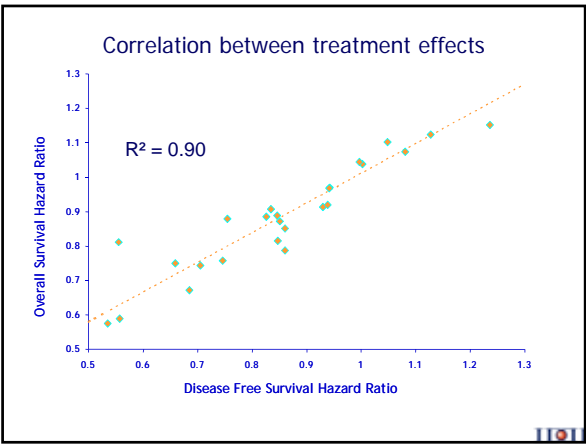
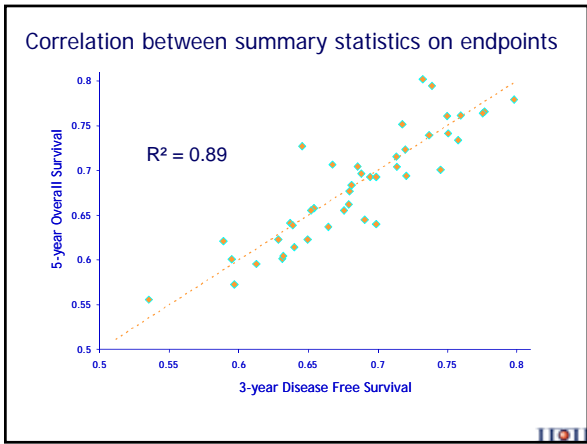
## Prediction of survival benefit



## Early colorectal cancer : DFS as a surrogate for survival

- 43 treatment arms in 18 randomized trials (20,898 patients)
  - 9 surgery alone control groups
  - 34 5FU-based experimental treatment groups
- Endpoints: disease-free survival and survival

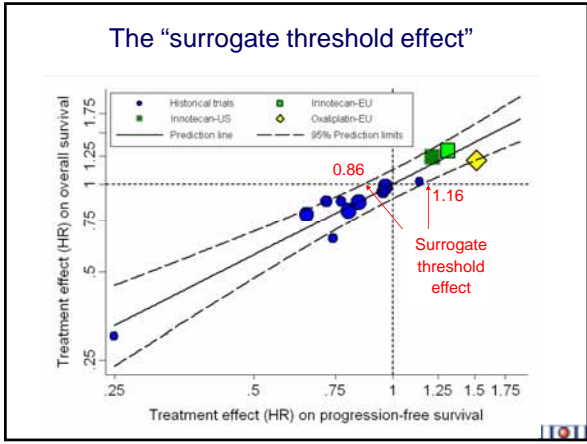
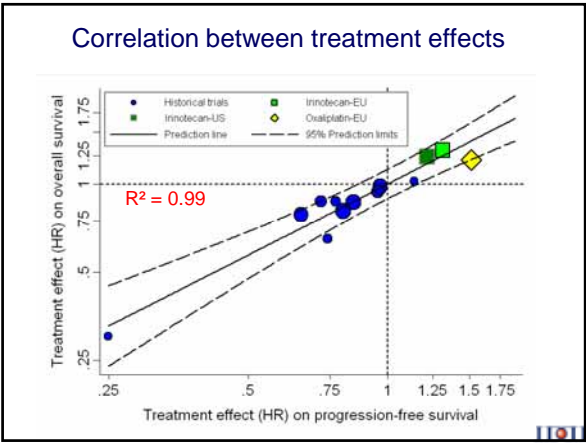
Ref: Sargent et al, JCO 2005;23:8664.



### PFS as a surrogate for survival: metastatic colorectal cancer

- 13 trials of first-line therapy on 4,352 patients
- Treatments (5FU + LV common arm):
  - Training set (10 trials, 3,089 patients)  
5FU Bolus or raltirexed vs. 5FU + LV
  - Validation set (3 trials, 1,263 patients)  
oxaliplatin or irinotecan + 5FU + LV vs. 5FU + LV

Ref: Buyse et al. J Clin Oncol 2007;25:5218.



### The “Surrogate Threshold Effect” (STE)

The “Surrogate Threshold Effect” is the treatment effect on the surrogate that would predict a statistically significant treatment effect on the true endpoint.

If STE is achievable, then the surrogate may be acceptable. In colorectal cancer,  $HR_{PFS} < 0.86$  have been achieved by experimental treatments.

A future trial could be planned in order to exclude the STE (i.e. the 95% confidence interval of the  $HR_{PFS}$  must lie entirely below  $STE = 0.86$ ).

Ref: Burzykowski and Buyse, Pharmaceutical Stat 2006, 5: 173.

### Issues with PFS as a surrogate for survival

- Evidence must come from meta-analyses, which include completed trials that do not address contemporary questions
- Surrogacy of PFS is highly tumor-dependent
- Surrogacy may also be treatment-dependent
- Should surrogacy be re-established for every new (class of) treatment(s)?



### Acknowledgments

Tomasz Burzykowski, Geert Molenberghs,  
Koji Oba, Xavier Paoletti, Pascal Piedbois

