# Methodological challenges


**Jan Bogaerts, PhD**

**Methodology Vice Director**

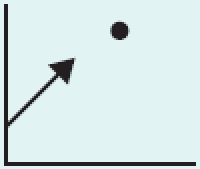
**EORTC**

# What is the question?

- The purpose
  - Regulatory approval
  - How to treat this particular patient?
  - What is the status of clinical understanding: how can we move ahead in this disease?

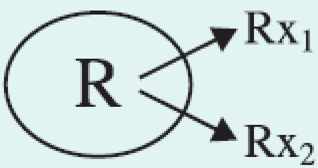
- The data

# Behera et al, Cancer Control 2007



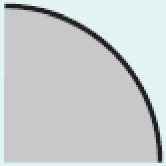
**A clear-enough future** (we can still be wrong but have a fairly clear picture what the future will look like)

- Treatment with dramatic side effects ("penicillin effect"). Non-RCTs appear to be optimal design to address this level of uncertainty.



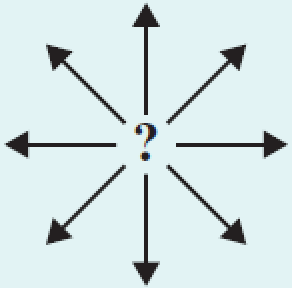

**Alternate futures** (a few discrete alternatives whose outcomes cannot be reliably predicted)

- Equipoise exists; an RCT the best method to resolve this level of uncertainty.



**A range of futures** (a range of potential futures can be identified but no natural discrete scenario has emerged)

- Many new drugs. Few data on safety and efficacy. Phase II trials to address this level of uncertainty.




**True ambiguity** (complete ignorance)

- A new chemical moiety. Further preclinical or phase I testing necessary to help shape our uncertainty in more solid direction.

# Levels of clinical evidence

| | |
|---|---|
| Level I | Adequately powered, high quality randomised trial, or meta-analysis of randomised trials showing statistically consistent results |
| Level II | Randomised trials inadequately powered, possibly biased, or showing statistically inconsistent results |
| Level III | Non-randomised studies with concurrent controls |
| Level IV | Non-randomised studies with historical controls (i.e. typical single arm phase II studies) |
| Level V | Expert committee review, case reports, retrospective studies |

*I. F. Tannock, Eur. J. Cancer Supplements Vol. 1, No. 5, Sept. 03, p. 93*

# (Part of) the Oxford 2011 Levels of Evidence

| Question | Step 1 (Level 1*) | Step 2 (Level 2*) | Step 3 (Level 3*) | Step 4 (Level 4*) |
|---|---|---|---|---|
| **How common is the problem?** | Local and current random sample surveys (or censuses) | Systematic review of surveys that allow matching to local circumstances** | Local non-random sample** | Case-series |
| **Is this diagnostic or monitoring test accurate?** (Diagnosis) | Systematic review of cross sectional studies with consistently applied reference standard and blinding | Individual cross sectional studies with consistently applied reference standard and blinding | Non-consecutive studies, or studies without consistently applied reference standards** | Case-contro "poor or nor reference st |
| **What will happen if we do not add a therapy?** (Prognosis) | Systematic review of inception cohort studies | Inception cohort studies | Cohort study or control arm of randomized trial* | Case-series control stud quality prog study** |
| **Does this intervention help?** (Treatment Benefits) | Systematic review of randomized trials or *n*-of-1 trials | Randomized trial or observational study with dramatic effect | Non-randomized controlled cohort/follow-up study** | Case-series, studies, or controlled st |
| **What are the COMMON harms?** (Treatment Harms) | Systematic review of randomized trials, systematic review of nested case-control studies, *n*-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect | Individual randomized trial or (exceptionally) observational study with dramatic effect | Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)** | Case-series, or historical studies** |
| **What are the RARE harms?** (Treatment Harms) | Systematic review of randomized trials or *n*-of-1 trial | Randomized trial or (exceptionally) observational study with dramatic effect | | |
| **Is this (early detection) test worthwhile?** (Screening) | Systematic review of randomized trials | Randomized trial | Non -randomized controlled cohort/follow-up study** | Case-series, or historical studies** |

# Oxford levels of evidence: tentative comments

- It is conceptually interesting to consider different questions
  - It does not seem to be very different from the ‘usual’ table
- Why would a systematic review of n-of-1 trials be better quality than a randomized trial ?
- One could consider amending/modifying this table with some ideas for rare cancers

- BUT …

# Levels of evidence for rare cancers

- **There is no logical rationale to say that the levels of evidence would function differently "because" it is harder to get data …**

# Considerations for design

- RCT remains the gold standard
- n-of-1 design: this is a sequence of different treatments in one and the same patient.
  - Has the feel of cross-over design
  - Question: how does that work in oncology?
- Play with the type I error (or even type II error). For example:
  - One sided testing: can be acceptable
  - Higher type I error (alpha): this will never be found, because the trial will not be repeated
  - More optimistic alternative hypothesis: this has the same practical effect as increasing the type II error (beta): only a really strong improvement has good chances of being identified. Look more at the confidence interval.

# Considerations for design

- Single-arm/non-comparative approaches
- The fact of having some responses is an improvement in itself
- The fact of stopping progression is an improvement in itself
- Robust historical data is available with small between trial variability (not likely, but happens)

# There is no current standard …

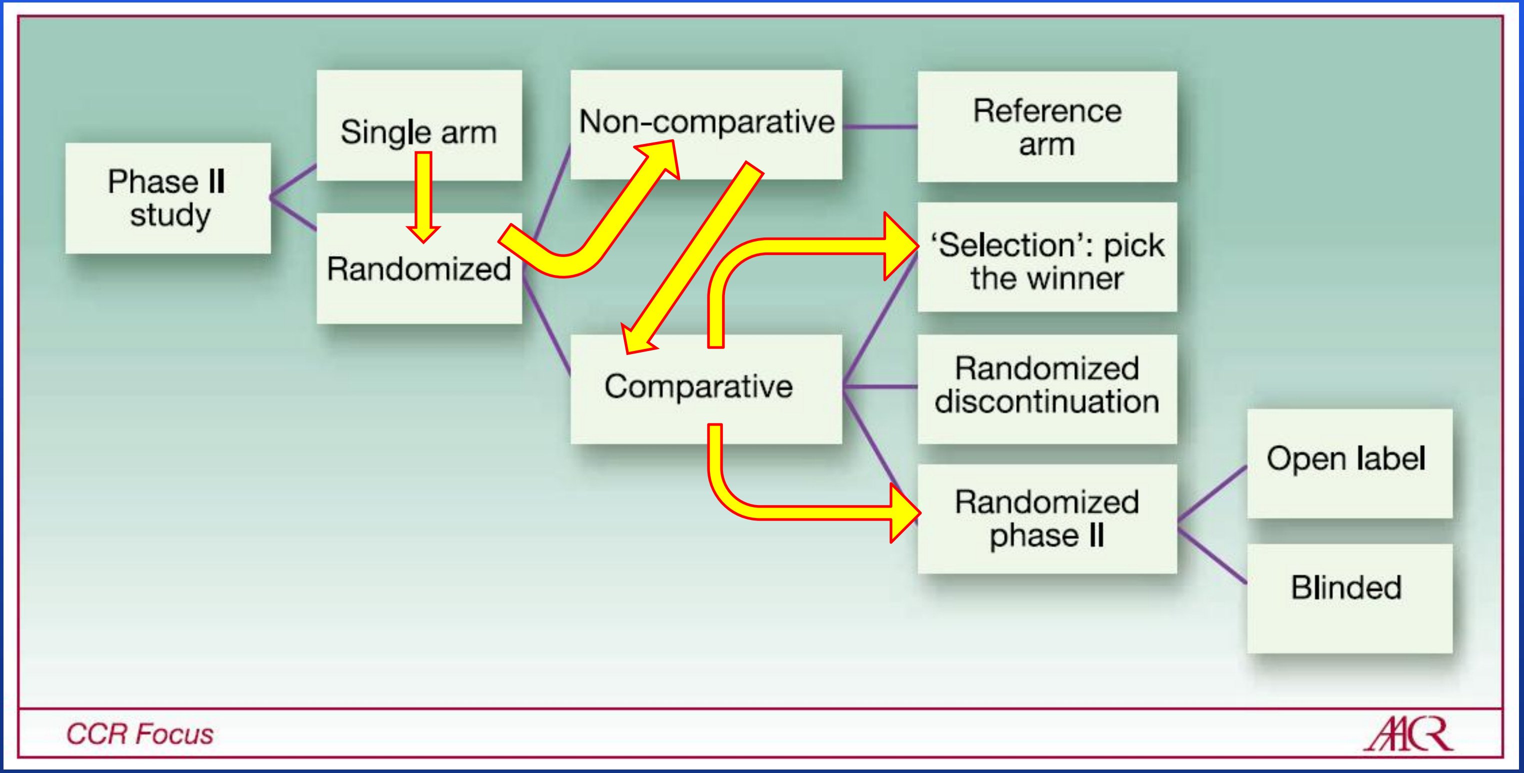
I would then (still) suggest a randomized approach with either a Phase II selection design, or a play-the-winner (adaptive randomization) approach

Other cases:

- Maybe it is worthwhile to incorporate in the plans a trial / decision point where disagreement is settled
- If the standard is wait-and-see, that can be randomized against

# Example of evolution: how we see Phase II trials

- Trying to improve the Positive Predictive Value
- Accommodate many objectives: moving to an amalgam of approaches




Seymour et al. CCR 2010

# Back-Up

# Looking for new common ground

- Trials with a high level of patient startup work
  - Screening many to obtain some eligible patients
  - Splitting according to markers
  - High workload to include patients
  - Timelines to enter a patient
- Think about:
  - Trials spanning several phases of development
  - Trials with multiple additional analyses / endpoints
  - TR analysis and planning of such analysis
  - Biobanking
  - Tools to perform complex logistics

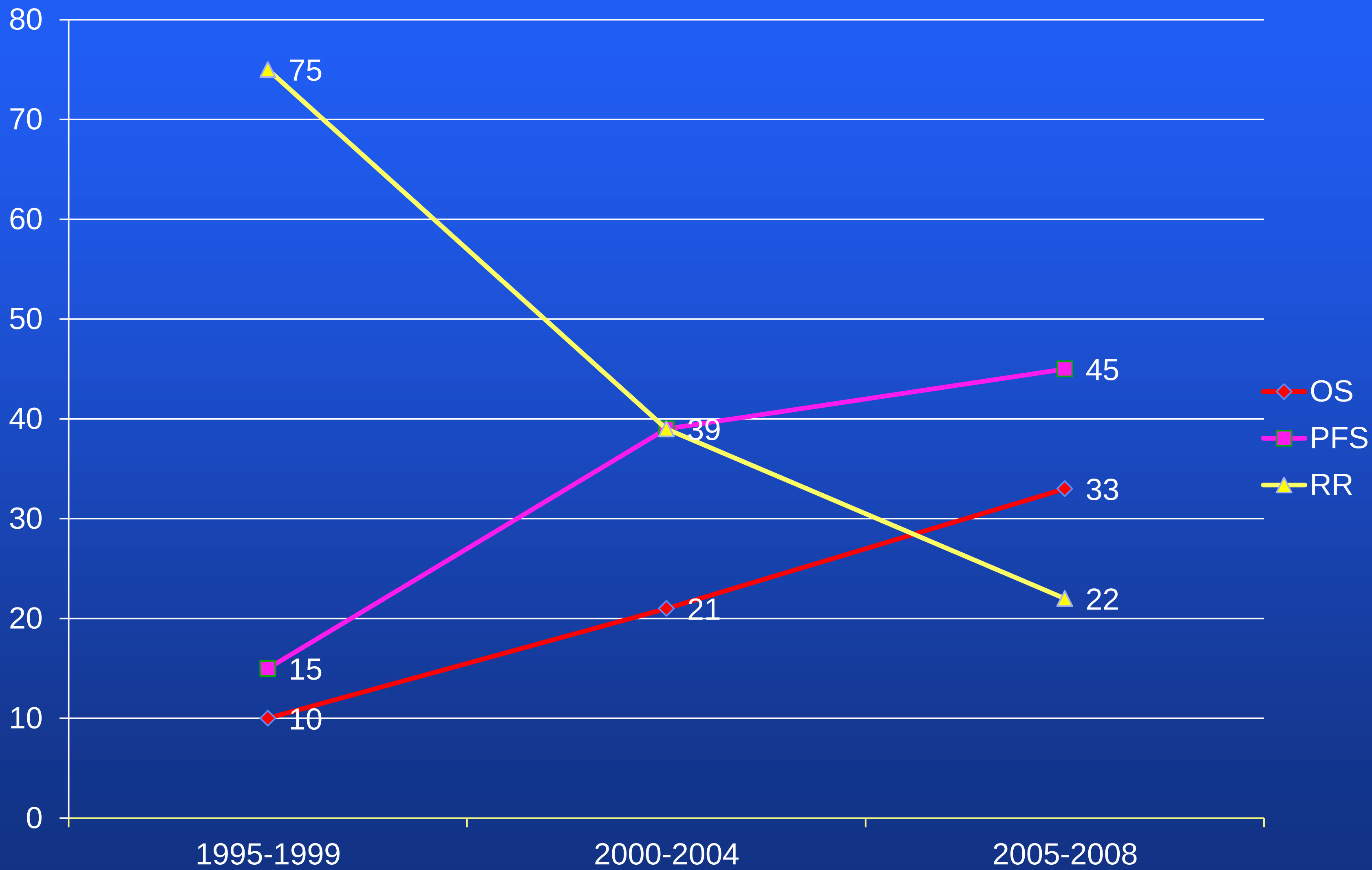
# Buzzword: Adaptive designs

- We are learning to plan and run these complicated trials in an acceptable way
  - Appropriate use of IDMC
  - Appropriate use of adaptive elements in the design
- Word of warning: adaptive designs are not the solution to manage the unexpected. But adaptive elements can be very interesting to manage the complicated.
- We are already using many adaptive ideas in our trials (all phases).
- Keys here are: think and discuss upfront and monitor during the trial

# FDA table of endpoints

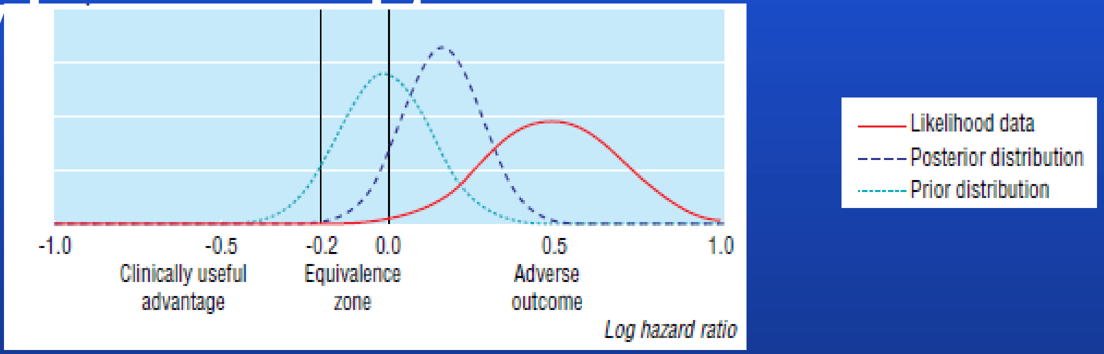
| Endpoint | Regulatory Evidence | Study Design | Advantages | Disadvantages |
|---|---|---|---|---|
| OS | Clinical benefit | Randomized | Direct measure of benefit, easy, precise | **Large studies, crossover / followup Tx affects**, noncancer deaths |
| Symptoms | Clinical benefit | Randomized, **blinded** | Patient perspective of direct clinical benefit | Blinding hard, missing data, clinically relevant effect, validated tools lacking |
| DFS | Surrogate | Randomized, **blinded**, **blinded review** | Smaller, shorter | **Not stat. validated as surrogate for OS / not precise, open to bias / many definitions** |
| RR | Surrogate | **Blinded, blinded review** | 1-arm possible, smaller, shorter, attributable to drug | No direct measure of benefit / no comprehensive measure of drug activity / only subset of benefiting pats. |
| CRR | Surrogate | **Blinded, blinded review** | 1-arm possible, smaller, shorter, durable CR = benefit | No direct measure of benefit / no comprehensive measure of drug activity / small subset of benefiting pats. |
| PFS | Surrogate | Randomized, **blinded**, **blinded review** | Smaller, shorter, SD included, crossover / other Tx not affecting, objective & quantitative | **Not stat. validated as surrogate for OS / not precise, open to bias /many definitions / frequent assessments / need to balance timing x arms** |

# Evolution of endpoints leading to EMA oncology approvals



# Alternative designs (cont'd)

- **Bayesian design, formally incorporating historical data into the design**
  - **Involve prior beliefs which may not be universally accepted**
  - **If we conduct a small trial, the choice of the prior may carry**




*Tan et al. Strategy for randomized clinical trials in rare cancers. BMJ, 2003*