Statistical Considerations in Preoperative Clinical Trials

Donald A. Berry
dberry@mdanderson.org

THE UNIVERSITY OF TEXAS MD ANDERSON CANCER CENTER
Dispense with sample size issue when pCR is primary endpoint:

Essentially same as for metastatic BC with tumor response as primary endpoint, and “interest in” PFS and OS
OUTLINE

- Are adjuvant trials still viable?
- Efficiency of neoadjuvant trials
- pCR as correlate or surrogate?
- Modeling pCR:DFS:OS
- Fine tuning pCR
CALGB node+ adjuvant trials

- CALGB 7581: N = 888
- CALGB 8082: N = 933
- CALGB 8541: N = 1550
- CALGB 9344: N = 3120

- Targeted # DFS events: 1800
- Interim analyses: 450, 900, 1350

Today!
ATAC: N=9366

Mean: 10 days

p=0.0013 for A vs T
Potential for more sensitive—and earlier!—comparisons in neoadjuvant trials: An example
Neoadjuvant Trastuzumab in HER2+ Breast Cancer*

\[ \text{FEC x 4 Taxol x 4} \]

\[ \text{FEC x 4 Taxol x 4 + trastuz} \]

\[ n = 82 \]

\[ \text{pCR?} \]

*Buzdar et al, *JCO* (2005)*
Data Monitoring Committee

- Annual monitoring by DMC
- Interim results after 34 patients:
  - Trastuzumab: 12/18 = 67%
  - Control: 4/16 = 25%
- Bayesian probability that outcome will still be significant after 164 patients: 95%
- ASCO —> JCO
Trastuzumab chronology

Metastatic  Buzdar  Adjuvant
1000s of pts  34 pts  1000s of pts

Neat link, though small
What about pCR?

● Great statistically because:
  ■ Fixed time of assessment
  ■ Early
  ■ Enables adaptive designs

● Should be fine tuned

● But is it a surrogate for anything of clinical relevance?
“Surrogate endpoint” (Prentice 1989)

- “a response variable for which a test of the null hypothesis of no relation to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint.”

- High hurdle: pCR doesn’t qualify

- But pCR is useful nonetheless!
Using neoadjuvant therapy in drug development: An adaptive example
Seamless phases II/III

- Primary breast cancer
- pCR *may* predict DFS, depending on treatment (not a “surrogate”)
- Primary endpoint: DFS
- Model pCR/DFS relationships
- Observe relationships—and “validate” within treatment group
"Standard" approach

DFS advantage
No DFS advantage

Market
Not

High pCR
Low pCR

Stop

Phase II
"White Space"
Phase III

12 mos
9-12 mos
> 5 yrs

Seamless phase II/III

< 4 yrs (usually)
Seamless phases

- Phase II: A few centers; 15 pts/mo, randomize equally to E vs C
- If predictive probs “look good,” expand (Phase III): Many centers; 60 pts/mo; initial centers continue accruing
- Max N = 1800

[Single trial: All data used in final analysis]
Early stopping

- Use pred probs of stat signif
- Frequent analyses (total of 18) using predictive probs to:
  - Switch to Phase III
  - Stop accrual for
    - Futility
    - Superiority
Comparisons

Conventional Phase III designs: Conv4 & Conv18, max N = 1800
(same significance level & power as adaptive Bayesian design)
Average N under $H_0$
Average $N$ under $H_1$
Advantages

- Duration of drug development shortened:
  - Fewer patients in trial
  - No hiatus for setting up phase III
  - All patients used for
    - Phase III endpoint
    - Relation between pCR & DFS

- N is seldom near 1800; when it is, it’s necessary!
Two reasons for advantages

• Exploiting pCR and its potential predictability
• Bayesian approach and frequent assessments of predictive probabilities
Further improvements possible in neoadjuvant settings (e.g., I-SPY2)

- Biomarkers
- Imaging
- Several drugs & combinations
- Adaptive randomization
OUTLINE

- Are adjuvant trials still viable?
- Efficiency of neoadjuvant trials
- pCR as correlate or surrogate?
- Modeling pCR:DFS:OS
- Fine tuning pCR