Surrogate Endpoints and Statistical Considerations

Donald Berry
<dberry@mdanderson.org>
Disclosure

Co-owner of Berry Consultants, LLC, a company that designs Bayesian adaptive clinical trials for pharmaceutical and medical device companies, NIH & NCI cooperative groups, patient advocacy groups, and international consortia.
Guidance for Industry
Pathological Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

October 2014
Clinical/Medical

IV. A. Rationale for Use of Pathological Complete Response as a Surrogate Endpoint in Neoadjuvant Trials
CTNeoBC, Cortazar et al.

Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis

Patricia Cortazar, Lijun Zhang, Michael Urrich, Keyur Mehta, Joseph P Costantino, Norman Wolmark, Hervé Bonnemoy, David Cameron, Jesus Correa, Pierre-Victor Delahaye, Graeme Le Feuvre, Tedd Tannock, Donald Edlich, Lothar Pissarek, and Rizwan Tongi

Findings

We obtained data from 12 identified international trials and 11,955 patients

Richard Pazdur, Nina Ditsch, Priya Rastogi, Wolfgang Hiermann, Gunter von Minckwitz

Summary

Pathological complete response has been proposed as a surrogate endpoint for prediction of long-term clinical benefit, such as disease-free survival, event-free survival (EFS), and overall survival (OS). We had four key objectives: to establish the association between pathological complete response and EFS and OS, to establish the definition of pathological complete response that correlates best with long-term outcome, to identify the breast cancer subtypes in which pathological complete response is best correlated with long-term outcome, and to assess whether an increase in frequency of pathological complete response between treatment groups predicts improved EFS and OS.

Methods

We searched PubMed, Embase, and Medline for clinical trials of neoadjuvant treatment of breast cancer. To be eligible, studies had to meet three inclusion criteria: include at least 200 patients with primary breast cancer treated with neoadjuvant chemotherapy followed by surgery; have available data for pathological complete response, EFS, and OS; and have a median follow-up of at least 3 years. We compared the three most commonly used definitions of pathological complete response—ypT0 ypN0, ypT0/IS ypN0, and ypT0/is— for their association with EFS and OS in a responder analysis. We assessed the association between pathological complete response and EFS and OS in various subgroups. Finally, we did a trial-level analysis to assess whether pathological complete response could be used as a surrogate endpoint for EFS or OS.

Findings

We obtained data from 12 identified international trials and 11,955 patients were included in our responder analysis. Eradication of tumor from both breast and lymph nodes (ypT0 ypN0 or ypT0/is ypN0) was better associated with benefit.
Cortazar patient-level analyses in Figure 5

**HER2-positive**

- Event-free survival (%)
- HR = 0.39

**Triple negative**

- Event-free survival (%)
- HR = 0.24

**Hormone-receptor-positive, HER2-negative**

- Event-free survival (%)
- HR = 0.49
Impressive! Why isn’t patient-level analyses enough for pCR to be a “validated” surrogate endpoint?
Comparing experimental and control arms for pCR vs non-pCR

Panel A: Exp same as control
Panel B: For exp, randomly selected non-pCR controls (20% of total) relabeled pCR
Panel C: For exp, non-pCR controls with non-events (20% of total) relabeled pCR

EFS for exp and cont identical within and across all 3 panels!
Improve pCR means improve EFS?
Cortazar trial-level analyses in Figure 6

Figure 6: Trial-level correlation between treatment effect on pathological complete response
Deficiencies and inefficiencies in Figure 6

- Ignores whether patients with longer EFS are those with pCRs
- Reduces information in 10,000 patients to 10 datapoints; loses 99.9% of information about correlation between EFS and pCR—even the sign of the correlation is difficult to estimate
- There is little treatment effect in the 10 RCTs; difficult to show correlation based on treatment effect when there is no treatment effect
- Requires RCTs
How else but Fig 6 for showing validated surrogacy?
Trade-off between pCR improvement and EFS by pCR effect

Panel D: Same as Panel A but pCR rate greater than control
Panel E: Experimental arm EFS increased by 50% for all non-pCR; pCR rate > 40%
Panel F: Same as Panel C [non-pCR controls with non-events (20% of total) relabeled pCR] but pCR rate > 60%
Predicting pCR rate, assuming panel A
Cont pCR rate 40% but with exp arm rate 60%

\[ HR = 0.19 \]
Predicting pCR rate, assuming panel A
Cont pCR rate 40% but with exp arm rate 60%

Sample size for statistical significance when HR = 0.86 for exp/cont: n=1800
The historical relationship between EFS and pCR may be different in a future trial.

So re-estimate the trial’s sample size adapting to the actual pCR rates by treatment, and the EFS by pCR relationships by treatment.
Does it work in practice?
HER2+ metaanalysis: Broglio et al. *JAMA Oncology* 2016

**Expected if ΔpCR predicts EFS perfectly**

HR=0.39
Summary

• Reasonably likely vs. validated surrogates
• Patient-level vs trial-level analyses
• Demonstrating surrogacy from RCTs
• Demonstrating surrogacy from single-arm trials
• Designing trials, learning about pCR/EFS