Phase II and III Immunotherapy Clinical Trials

ELIZABETH GARRETT-MAYER, PHD
DIVISION DIRECTOR, BIOSTATISTICS AND RESEARCH DATA GOVERNANCE
CENTER FOR RESEARCH AND ANALYTICS (CENTRA)
AMERICAN SOCIETY OF CLINICAL ONCOLOGY
Disclosures

Views and ideas are my own and do not represent American Society of Clinical Oncology

I am the primary statistician for the TAPUR study, a precision medicine phase II basket trial with 17 different treatment regimens, including “Nivolumab and Ipilimumab” and “Pembrolizumab”
Topics

Levels of Evidence

Endpoints

Trial types
- Single Arm Phase II trials
- Randomized Phase II Trials
- Randomized Phase III Trials

Data display and interpretation

Challenges and Considerations for Immunotherapy Trials
- Interpreting trials with cross-over
- Comparing to traditional therapies
Levels of Evidence of Efficacy

Low
- Single arm Phase II Trial

Medium
- Randomized Phase II Trial

High
- Randomized Phase III Trial
What makes a cancer therapy “efficacious”?  

It delays time to death, compared to other available treatments  

It maintains or improves quality of life  

Assumptions regarding tumor burden:  
- Shrinking tumor burden **should** lead to longer survival  
- Delayed progression **should** lead to longer survival  

Minimizing toxicities (adverse events, especially serious ones) is important
**Objective Response (OR):**
- Partial or complete tumor response
- Significant shrinkage of ‘target lesions’ and no new lesions.
- **RECIST criteria:** Response Evaluation Criteria in Solid Tumors
- **ir-RECIST criteria:** Immune-related Response Evaluation Criteria in Solid Tumors

**Progression-Free Survival (PFS, or ir-PFS):**
- Time from treatment initiation (or randomization) until PROGRESSION or DEATH from any cause
- Progression based on **RECIST** or **ir-RECIST**

**Overall Survival (OS):**
- Time from treatment initiation (or randomization) until DEATH

---

**Surrogate measure**

**Gold Standard**

Strength

- Low
- Medium
- High
Common Endpoint in Cancer Immunotherapy Trials

Duration of Response (DOR)
- Time from Objective Response (OR) to disease progression
- Can only be measured in patients who have an objective response

Why DOR in cancer immunotherapy trials?
- Some patients have an exceptional response with long duration.
- Different than in chemotherapy which tends to delay progression in patients with metastatic disease
- Look for ‘swimmer plots’ and ‘spider plots’
Swimmer Plot

Spider Plot

Single Arm Phase II

**Follows Phase I, initial look at efficacy**

**Enroll all patients on a SINGLE treatment arm**
- Sample size usually around 20 to 50 patients
- Common endpoint is **Objective Response** (RECIST or ir-RECIST)
- Can be combination therapy

**Common when the target patient population is relatively RARE**
- Biomarker required for eligibility
- Rare cancer

**Benefits**
- Relatively small sample size

**Limitation**
- Without a comparator arm, difficult to conclude “success”
Primary objective: To estimate the proportion of patients who have a brain metastasis response.

Endpoint: brain metastasis response

Sample size: Target N = 44 (actual N = 37)

25% response rate considered “success”

Results: 30% response rate observed
Pembrolizumab for management of patients with NSCLC and brain metastases: long-term results and biomarker analysis from a non-randomised, open-label, phase 2 trial

Lancet Oncology, May 2020, 21(5):655-663

**Primary objective:** To estimate the proportion of patients who have a brain metastasis response.

**Endpoint:** brain metastasis response

**Sample size:** Target N = 44 (actual N = 37)

25% response rate considered “success”

**Results:** 30% response rate observed

Swimmer plot including patients who had a brain metastasis response or remained on trial for at least 4 months (19 patients).
Primary objective: To estimate the proportion of patients who have a brain metastasis response.

Endpoint: brain metastasis response

Sample size: Target N = 44 (actual N = 37)

25% response rate considered “success”

Results: 30% response rate observed
Phase II Basket Trials

Precision medicine

Tests how well a new drug or other substance works in patients who have different types of cancer that all have the same mutation or biomarker. (NCI, www.cancer.gov)

Focus is more on the genetic/genomic make-up of the tumor than on the site of the tumor

Example: TAPUR (Targeted Agent and Profiling Utilization Registry) Trial

Figure 1:
Participant registration, enrollment and cohort assignment process. Panel A displays the process by which participants are registered and enrolled into the study. Panel B displays the organization of the cohorts for analysis which are grouped by treatment, targeted variant and tumor type.

Mangat, Halabi, Bruinooge et al., JCO Precision Oncology, 2018
Randomized Phase II Trials

At least two treatment groups

Patients are randomly assigned to groups
  - Might be masked (i.e., arm assignment is unknown)

Comparison arms can take different forms
  - Combination vs. single agent
  - Two different doses or schedules of same treatment
  - Experimental agents vs. standard of care

What makes it phase II vs. phase III?
  - Endpoint choice
  - Sample size (~50 to several hundred patients)

Phase III usually still required after a “successful” randomized phase II trial.
Randomized Phase II example

Annals of Oncology. 2013 Jan;24(1):75-83

**Design:** Patients with chemotherapy-naïve ED-SCLC were randomized 1: 1: 1 to receive paclitaxel/carboplatin with either (A) placebo or (B) concurrent ipilimumab or (C) phased ipilimumab

**Objective:** Compare ir-PFS in ipilimumab groups vs. placebo

**Sample size:** Target N = 130 (~43 per group)

**Primary endpoint:** ir-PRS
Randomized Phase

**Design:** Patients with chemotherapy-naïve ED-SCLC were randomized 1:1:1 to receive paclitaxel/carboplatin with either (A) placebo or (B) concurrent ipilimumab or (C) phased ipilimumab

**Objective:** Compare ir-PFS in ipilimumab groups vs. placebo

**Sample size:** Target N = 130 (~43 per group)

**Primary endpoint:** ir-PRS

**Result:** ir-PFS in phased ipilimumab has longer ir-PFS than placebo

*Annals of Oncology. 2013 Jan;24(1):75-83*
Considerations for Trial Designs

**Evaluability of patients:**
- Patients who leave the study for reasons unrelated to treatment or disease?
- Patients who enroll but receive no therapy or just a small amount (< 1 cycle) of treatment?

**Timing of measurements**
- When OR or PFS is endpoint, how often to assess disease?
- Needs to be consistent with other trials in same population
- Should be convenient for patients (i.e., time it with treatment visit)

**Quality of Life and/or Patient Reported Outcomes**
- Important to ensure patient well-being is captured, assessed, compared.
Challenges in Immunotherapy Trials

Pseudo-progression
- “Pseudo-progression is a phenomenon in which an initial increase in tumor size is observed or new lesions appear, followed by a decrease in tumor burden; this phenomenon can benefit patients receiving immunotherapy but often leads to premature discontinuation of treatment owing to the false judgment of progression.”
- Use ir-RECIST to help mitigate issue

Delayed responses
- Different than cytotoxics
- Challenging for adaptive trial designs using OR as endpoint

Non-specific or heterogeneous adverse event (AE) profile
- Traditional anti-cancer agents have predictable and/or consistent toxicities
- Immunotherapies affect patients in various ways.
  - Attribution of AEs affected
  - Patterns of AEs harder to discern
Randomized Phase III Trials

Similar to randomized phase II, but designs include:
- More relevant endpoint
- Larger sample size
- Inferences are more definitive; less exploratory

Often have overall survival as the primary endpoint
- More challenging as more treatment options are available
- “Cross-over” can confound inferences

“Powered” to detect a clinically meaningful difference
- That is, sample size is sufficiently large.

Designed to change treatment paradigm
- Limited comparisons considered
- Usually, experimental regimen vs. standard of care
Design: Patient with advanced melanoma randomized (2:1) nivolumab vs. investigator’s choice chemotherapy

Objective: Compare OS in two treatment arms

Sample size: N = 405

Primary endpoint: Overall survival

Result: Higher, more durable responses with longer DOR in nivolumab arm, but no difference in OS
Swimmer plot
(remember—2:1 randomization)

27% response rate in Nivo (N=74)
10% response rate in ICC (N=13)
Issues with Cross-Over

Upon progression, patients will receive another treatment (maybe ICI).

“Given the higher number of ICC patients who received subsequent systemic treatment, OS was investigated in a sensitivity analysis by censoring at the start of the PD-1/PD-L1 therapy that was received after assigned therapy in the ICC group.”

Cross-over within protocol: Ethical approach, encourages enrollment.

However, groups become:
- Nivolumab, or Nivolumab followed by ICI
- ICC, or ICC followed by ICI (or something else)

What happens if we ‘censor’ death times for patients who cross-over to Nivolumab?
- Looks like Nivolumab has better survival
- But...selection bias!
“Long tail”

Common measure of overall treatment effect is the **hazard ratio**.

Assumes “proportional risk” of events over time.

**Shapes** of the Kaplan-Meier curves for traditional agents and immunotherapies are different:
- Proportionality is violated
- Hazard ratio is not valid

New measures are needed to quantify treatment effect which has multiple components.