HERCULES: a 3-arm noninferiority trial in advanced, HER2-positive cancers of the stomach and gastroesophageal junction

Ethan Ashby[1], *Deepa Oja*[2], *Brenda Osei-Assibey*[2], *Makayla Tang*[1], *Mihkai Wickline*[2]

[1] ***University of Washington, Department of Biostatistics***

[2] ***University of Washington, School of Nursing***

Biostat524 Site Visit
Tuesday 7th June, 2022

Contents

Where we were; where we are

Where we were: $\sim$ 3 weeks ago, our protocol outlined a 2-arm superiority trial comparing HER-Vaxx plus chemotherapy vs. trastuzumab plus chemotherapy at NCI clinics in the United States.

Where we were; where we are

Where we were: $\sim$ 3 weeks ago, our protocol outlined a 2-arm superiority trial comparing HER-Vaxx plus chemotherapy vs. trastuzumab plus chemotherapy at NCI clinics in the United States.

- We had concerns regarding whether HER-Vaxx could reasonably improve overall survival compared to trastuzumab.

Where we were; where we are

Where we were: $\sim$ 3 weeks ago, our protocol outlined a 2-arm superiority trial comparing HER-Vaxx plus chemotherapy vs. trastuzumab plus chemotherapy at NCI clinics in the United States.

- We had concerns regarding whether HER-Vaxx could reasonably improve overall survival compared to trastuzumab.
- We had practical and ethical reservations regarding conducting a trial in a country with relatively low disease burden.

Where we were; where we are

Where we were: $\sim$ 3 weeks ago, our protocol outlined a 2-arm superiority trial comparing HER-Vaxx plus chemotherapy vs. trastuzumab plus chemotherapy at NCI clinics in the United States.

- We had concerns regarding whether HER-Vaxx could reasonably improve overall survival compared to trastuzumab.
- We had practical and ethical reservations regarding conducting a trial in a country with relatively low disease burden.
- Due to limited global access to trastuzumab treatment, we believed that HER-Vaxx held great potential from a global oncology perspective.

Where we were; where we are

Where we are: our concerns led us to reformulate our trial as a non-inferiority trial comparing HER-Vaxx plus chemotherapy and trastuzumab plus chemotherapy.

Where we were; where we are

Where we are: our concerns led us to reformulate our trial as a non-inferiority trial comparing HER-Vaxx plus chemotherapy and trastuzumab plus chemotherapy.

- Dubious constancy assumption in the pivotal ToGA trial (2010) $\implies$ 3-arm non-inferiority design.

Where we were; where we are

Where we are: our concerns led us to reformulate our trial as a non-inferiority trial comparing HER-Vaxx plus chemotherapy and trastuzumab plus chemotherapy.

- Dubious constancy assumption in the pivotal ToGA trial (2010) $\implies$ 3-arm non-inferiority design.
- Needed more participants and unethical to enroll a "chemotherapy only" arm in a U.S. trial $\implies$ relocated to China where disease burden is high and access to anti-HER2 therapy is limited.

MANI trials

- Enroll placebo, active control, and experimental arms.

MANI trials

- Enroll placebo, active control, and experimental arms.
- Inclusion of placebo arm abrogates constancy assumption, since active control effect versus placebo is directly measured.

MANI trials

- Enroll placebo, active control, and experimental arms.
- Inclusion of placebo arm abrogates constancy assumption, since active control effect versus placebo is directly measured.
- If experimental and active control treatments are not shown superior to placebo, this suggests the trial lacks assay sensitivity – a hypothesis we can test to avoid claiming non-inferiority under "similarly ineffective" paradigm.

MANI trials

- Enroll placebo, active control, and experimental arms.
- Inclusion of placebo arm abrogates constancy assumption, since active control effect versus placebo is directly measured.
- If experimental and active control treatments are not shown superior to placebo, this suggests the trial lacks assay sensitivity – a hypothesis we can test to avoid claiming non-inferiority under "similarly ineffective" paradigm.
- Our mission: in order for HER-Vaxx plus chemotherapy to be non-inferior to trastuzumab plus chemotherapy it must (a) be superior to chemotherapy alone and (b) retain a clinically-relevant proportion of trastuzumab's effect.

MANI trials

- Enroll placebo, active control, and experimental arms.
- Inclusion of placebo arm abrogates constancy assumption, since active control effect versus placebo is directly measured.
- If experimental and active control treatments are not shown superior to placebo, this suggests the trial lacks assay sensitivity – a hypothesis we can test to avoid claiming non-inferiority under "similarly ineffective" paradigm.
- Our mission: in order for HER-Vaxx plus chemotherapy to be non-inferior to trastuzumab plus chemotherapy it must (a) be superior to chemotherapy alone and (b) retain a clinically-relevant proportion of trastuzumab's effect.
- Equivalent to testing $H_0^{NI} : H_0^{AS} \cup H_0^{RE}$ versus $H_a^{NI} : H_a^{AS} \cap H_a^{RE}$ $\implies$ testing sub-hypotheses at level α ensures global hypothesis is level α.

Inequity in global oncology

The landscape of global oncology is inequitable.

- Nearly 89% of cancer therapies launched in 2011-2017 were available in the U.S. in 2018, while only 8% were available in China (PhRMA, 2018).

Inequity in global oncology

The landscape of global oncology is inequitable.

- Nearly 89% of cancer therapies launched in 2011-2017 were available in the U.S. in 2018, while only 8% were available in China (PhRMA, 2018).
- Survey of HER2 testing for gastric cancer in China estimated 73% of cancers do not receive testing (Chambers *et al.* 2020)

Inequity in global oncology

The landscape of global oncology is inequitable.

- Nearly 89% of cancer therapies launched in 2011-2017 were available in the U.S. in 2018, while only 8% were available in China (PhRMA, 2018).
- Survey of HER2 testing for gastric cancer in China estimated 73% of cancers do not receive testing (Chambers *et al.* 2020)
- Trastuzumab treatment for confirmed HER2-positive breast cancer is 10-50% in China, likely lower for gastric cancers.

Inequity in global oncology

The landscape of global oncology is inequitable.

- Nearly 89% of cancer therapies launched in 2011-2017 were available in the U.S. in 2018, while only 8% were available in China (PhRMA, 2018).
- Survey of HER2 testing for gastric cancer in China estimated 73% of cancers do not receive testing (Chambers *et al.* 2020)
- Trastuzumab treatment for confirmed HER2-positive breast cancer is 10-50% in China, likely lower for gastric cancers.
- Neighboring countries in East Asian countries – Mongolia, Bhutan – the highest gastric cancer ASRs in the world but virtually no centralized cancer care.

HER-Vaxx in resource-limited settings

- Cost
  - Inflation-adjusted cost of adding trastuzumab treatment to SoC in East Asia is $> \$23,000$ per patient (Shiroiwa 2011).
  - Cost of peptide vaccines are as low as $\$5 - \$10$/dose. 10-dose HER-Vaxx regimen could yield 10s-100s-fold savings.

HER-Vaxx in resource-limited settings

- Cost
  - Inflation-adjusted cost of adding trastuzumab treatment to SoC in East Asia is $> \$23,000$ per patient (Shiroiwa 2011).
  - Cost of peptide vaccines are as low as $\$5 - \$10$/dose. 10-dose HER-Vaxx regimen could yield 10s-100s-fold savings.
- Ease of administration
  - Trastuzumab IV infusion - more time-consuming, resource-intensive, and costly than IM injection.
  - Trastuzumab requires consistent bi/tri-weekly redosing, whereas HER-Vaxx may require fewer treatment visits.

HER-Vaxx in resource-limited settings

- Cost
  - Inflation-adjusted cost of adding trastuzumab treatment to SoC in East Asia is $> \$23,000$ per patient (Shiroiwa 2011).
  - Cost of peptide vaccines are as low as $\$5 - \$10$/dose. 10-dose HER-Vaxx regimen could yield 10s-100s-fold savings.
- Ease of administration
  - Trastuzumab IV infusion - more time-consuming, resource-intensive, and costly than IM injection.
  - Trastuzumab requires consistent bi/tri-weekly redosing, whereas HER-Vaxx may require fewer treatment visits.

HER-Vaxx in resource-limited settings

- Cost
  - Inflation-adjusted cost of adding trastuzumab treatment to SoC in East Asia is $> \$23,000$ per patient (Shiroiwa 2011).
  - Cost of peptide vaccines are as low as $\$5 - \$10$/dose. 10-dose HER-Vaxx regimen could yield 10s-100s-fold savings.
- Ease of administration
  - Trastuzumab IV infusion - more time-consuming, resource-intensive, and costly than IM injection.
  - Trastuzumab requires consistent bi/tri-weekly redosing, whereas HER-Vaxx may require fewer treatment visits.
- Safety and side effects
  - HER-Vaxx may be associated with lower risks of adverse infusion reactions (anaphylaxis) and cardiotoxicity.

HER-Vaxx in resource-limited settings

- Cost
  - Inflation-adjusted cost of adding trastuzumab treatment to SoC in East Asia is $> \$23,000$ per patient (Shiroiwa 2011).
  - Cost of peptide vaccines are as low as $\$5 - \$10$/dose. 10-dose HER-Vaxx regimen could yield 10s-100s-fold savings.
- Ease of administration
  - Trastuzumab IV infusion - more time-consuming, resource-intensive, and costly than IM injection.
  - Trastuzumab requires consistent bi/tri-weekly redosing, whereas HER-Vaxx may require fewer treatment visits.
- Safety and side effects
  - HER-Vaxx may be associated with lower risks of adverse infusion reactions (anaphylaxis) and cardiotoxicity.
- Potentially enhanced mechanism of action
  - Multi-epitope formulation may offer more complete blockade of HER2 pathway, delaying progression/resistance/recurrence.
  - HER-Vaxx may extend survival *post*-progression.

Backup slides

Key trials we reference

- ToGA (2010): compared addition of trastuzumab to chemo for HER2-positive gastric/GEJ cancers. Participants not excluded on basis of HER2 expression level. Randomized 600 patients over 3 yrs. Addition of trastuzumab produced a 2.7 month improvement in OS, but benefit was restricted to the ICH3+/ICH2+ & FISH+ subgroup (4.2 month increase in OS, HR=0.65).
- JACOB (2018): compared addition of pertuzumab (another anti-HER2 mAb) to trastuzumab & chemo in advanced ICH3+/ICH2+ & FISH+ gastric/GEJ cancers. Randomized 800 participants over 2.5 year period between (2013-2016). Addition of pertuzumab led to non-significant 3.3 mo. increase in OS (p=0.057).
- EVIDENCE (2021): open-label registry trial to assess real-world treatment patterns, safety, and efficacy of trastuzumab for gastric cancer in China. Enrolled 1600 participants over 2.75-year span (2013-2016). Trastuzumab treatment was associated with better propensity-score matched survival outcomes & better tumor responses.

MANI trials

Traditional two-arm non-inferiority designs require the constancy assumption to achieve valid results.

MANI trials

Traditional two-arm non-inferiority designs require the constancy assumption to achieve valid results.

A trial must also have assay sensitivity, meaning the the trial is capable of detecting differences between regimens. Assay sensitivity can be influenced by a variety of factors including statistical power, adherence to treatment, proper disease diagnosis, etc.

MANI trials

Traditional two-arm non-inferiority designs require the constancy assumption to achieve valid results.

A trial must also have assay sensitivity, meaning the the trial is capable of detecting differences between regimens. Assay sensitivity can be influenced by a variety of factors including statistical power, adherence to treatment, proper disease diagnosis, etc.

Multi-arm non-inferiority trials (MANI) are an appealing alternative to 2-arm trials.

MANI trials

Traditional two-arm non-inferiority designs require the constancy assumption to achieve valid results.

A trial must also have assay sensitivity, meaning the the trial is capable of detecting differences between regimens. Assay sensitivity can be influenced by a variety of factors including statistical power, adherence to treatment, proper disease diagnosis, etc.

Multi-arm non-inferiority trials (MANI) are an appealing alternative to 2-arm trials.

1. Contemporaneous control group abrogates the need for constancy assumption.

MANI trials

Traditional two-arm non-inferiority designs require the constancy assumption to achieve valid results.

A trial must also have assay sensitivity, meaning the the trial is capable of detecting differences between regimens. Assay sensitivity can be influenced by a variety of factors including statistical power, adherence to treatment, proper disease diagnosis, etc.

Multi-arm non-inferiority trials (MANI) are an appealing alternative to 2-arm trials.

1. Contemporaneous control group abrogates the need for constancy assumption.
2. Assay sensitivity is now testable.

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?
   - We propose retention of at least 60% of trastuzumab's effect to claim non-inferiority, corresponding to an estimated improvement in median OS from 10.85 to 13.25 months (2.4 months).

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?
   - We propose retention of at least 60% of trastuzumab's effect to claim non-inferiority, corresponding to an estimated improvement in median OS from 10.85 to 13.25 months (2.4 months).
3. How many patients and in what allocation do we need?

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?
   - We propose retention of at least 60% of trastuzumab's effect to claim non-inferiority, corresponding to an estimated improvement in median OS from 10.85 to 13.25 months (2.4 months).
3. How many patients and in what allocation do we need?
   - Mielke, Munk, and Schact (2008) provide a framework for allocation, sample size calculation, and inference.

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?
   - We propose retention of at least 60% of trastuzumab's effect to claim non-inferiority, corresponding to an estimated improvement in median OS from 10.85 to 13.25 months (2.4 months).
3. How many patients and in what allocation do we need?
   - Mielke, Munk, and Schact (2008) provide a framework for allocation, sample size calculation, and inference.
4. What interim monitoring procedures do we use?

Considerations when designing MANI trials

1. Is enrolling a contemporaneous control group ethical?
   - We argue yes, but *only* in a setting where the active control is not standard of care.
2. What is the retention of effect target?
   - We propose retention of at least 60% of trastuzumab's effect to claim non-inferiority, corresponding to an estimated improvement in median OS from 10.85 to 13.25 months (2.4 months).
3. How many patients and in what allocation do we need?
   - Mielke, Munk, and Schact (2008) provide a framework for allocation, sample size calculation, and inference.
4. What interim monitoring procedures do we use?
   - Interim testing and stopping procedures developed by Ochiai *et al.* 2017

Justifying NI margin

We chose NI margin: retention of 60% of the effect of the active control, corresponding to a clinically-relevant minimum 2.4 month improvement in OS under our sample size parameters.

Justifying NI margin

We chose NI margin: retention of 60% of the effect of the active control, corresponding to a clinically-relevant minimum 2.4 month improvement in OS under our sample size parameters.

We chose to define the NI margin as a *fraction of the assay sensitivity* rather than a *fixed margin* for a few reasons:

Justifying NI margin

We chose NI margin: retention of 60% of the effect of the active control, corresponding to a clinically-relevant minimum 2.4 month improvement in OS under our sample size parameters.

We chose to define the NI margin as a *fraction of the assay sensitivity* rather than a *fixed margin* for a few reasons:

1. Uncertain est. effects $\implies$ assigning fixed margin difficult
   - 1.1 e.g., 1-month shorter survival may be acceptable for a treatment that extends life by 12 months, but not for a treatment that extends survival by 2 months.

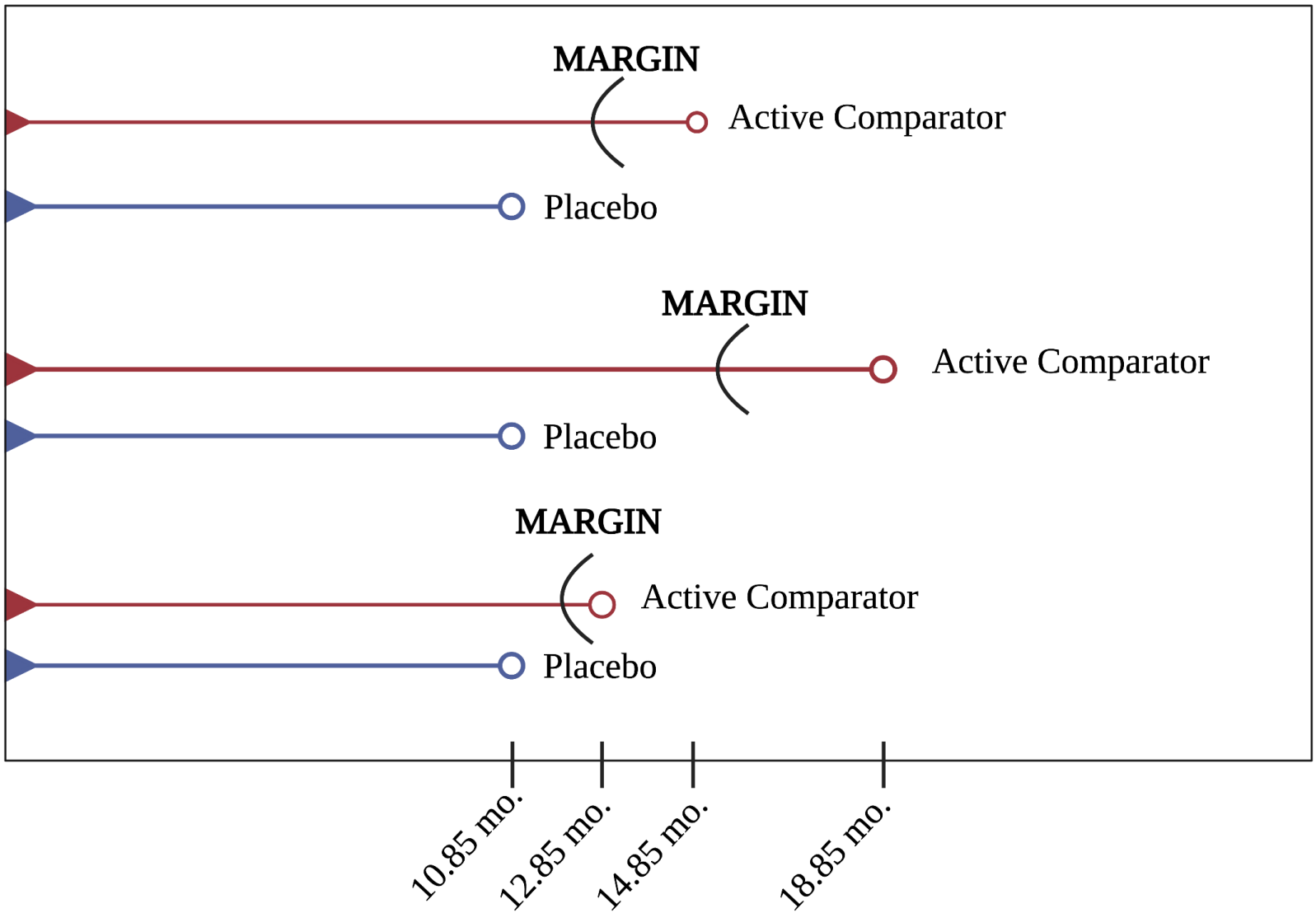
Justifying NI margin

We chose NI margin: retention of 60% of the effect of the active control, corresponding to a clinically-relevant minimum 2.4 month improvement in OS under our sample size parameters.

We chose to define the NI margin as a *fraction of the assay sensitivity* rather than a *fixed margin* for a few reasons:

1. Uncertain est. effects $\implies$ assigning fixed margin difficult
   - 1.1 e.g., 1-month shorter survival may be acceptable for a treatment that extends life by 12 months, but not for a treatment that extends survival by 2 months.
2. If HER-Vaxx was to *replace trastuzumab* in clinical care, one would need to demonstrate that HER-Vaxx is non-inferior in *months survived* (fixed margin). As a first step, we hope HER-Vaxx will *replace chemotherapy alone* and retain a substantial fraction of trastuzumab's effect (fraction margin).
   - 2.1 If HER-Vaxx is shown non-inferior in this trial, we may conduct a second non-inferiority trial with a fixed margin for non-inferiority.

Justifying NI margin



Interim Monitoring

Two goals: (a) monitor the trial for efficacy/futility and (b) adaptive trial design to ensure high power when initial design assumptions do not hold.

Interim Monitoring

Two goals: (a) monitor the trial for efficacy/futility and (b) adaptive trial design to ensure high power when initial design assumptions do not hold.

We propose using the DF-A stopping rule (Ochiai *et al.* 2017) for interim testing of $H_0 : H_0^{a.s.} \cup H_0^{r.e.}$. Assay sensitivity hypothesis $H_0^{a.s.}$ tested first, then tests retention of effect $H_0^{r.e.}$ once assay sensitivity is demonstrated. Multiplicity adjustment using O'Brien-Fleming boundary.

AMENDMENT: we propose a single sample size recalculation at 50% information using conditional power calculation based on CHW statistics as described by Ochiai *et al.* (2017).