NISS-Merck Meet-Up
Webinar
10 September 2019

To Be Or Not To Be
(A Subgroup)

Stephen J. Ruberg, PhD
President
Analytix Thinking, LLC
AnalytixThinking@gmail.com
AnalytixThinking.blog
Acknowledgement

My thinking on this problem has been significantly influenced by former colleagues at Lilly, most notably

Lei Shen
Rick Higgs
Ilya Lipkovich
Alice in Wonderland
Alice: "Would you tell me, please, which way I ought to go from here?"
Alice in Wonderland

**Alice:** "Would you tell me, please, which way I ought to go from here?"

**Cheshire Cat:** "That depends a good deal on where you want to get to."
Alice: "Would you tell me, please, which way I ought to go from here?"

Cheshire Cat: "That depends a good deal on where you want to get to."

Alice: "I don't much care where—"
Alice in Wonderland

Alice: "Would you tell me, please, which way I ought to go from here?"

Cheshire Cat: "That depends a good deal on where you want to get to."

Alice: "I don't much care where—"

Cheshire Cat: "Then it doesn't matter which way you go."
Alice: "Would you tell me, please, which way I ought to go from here?"

Cheshire Cat: "That depends a good deal on where you want to get to."

Alice: "I don't much care where—"

Cheshire Cat: "Then it doesn't matter which way you go."

Alice: "—so long as I get SOMEWHERE."
Alice in Wonderland

Alice: "Would you tell me, please, which way I ought to go from here?"

Cheshire Cat: "That depends a good deal on where you want to get to."

Alice: "I don't much care where—"

Cheshire Cat: "Then it doesn't matter which way you go."

Alice: "—so long as I get SOMEWHERE."

Cheshire Cat: "Oh, you're sure to do that, if you only walk long enough."
Key Messages

1. Be clear on your terminology
2. Be clear on your goals
3. Disciplined subgroup search (DSS)
4. Always do subgroup analysis identification
5. One last idea
1. Some Terminology
1. Some Terminology

Precision medicine
Targeted therapy/therapeutics

vs

Personalized medicine

Targeted therapeutics = *patients like YOU*

Personalized medicine = YOU
1. Some Terminology

Precision medicine
Targeted therapy/therapeutics
vs
Personalized medicine

Targeted therapeutics = *patients like YOU*

Personalized medicine = YOU
A targeted therapeutic allows the sponsor to make a regulatory approved claim of an expected treatment effect (efficacy or safety):

✓ Where the indicated patient population is itself selected by a biochemical, genetic or imaging biomarker.

✓ Where a subgroup of the indicated patient population (selected based on a clinical, biochemical, genetic, or imaging biomarker) is expected to have a differential treatment effect.

✓ Where individualized dosing (dose and/or duration) is based on a biomarker responsive to treatment.
1. Some Terminology

A targeted therapeutic allows the sponsor to make a regulatory approved claim of an expected treatment effect (efficacy or safety):

✓ Where the indicated patient population is itself selected by a biochemical, genetic or imaging biomarker.

✓ Where a subgroup of the indicated patient population (selected based on a clinical, biochemical, genetic, or imaging biomarker) is expected to have a differential treatment effect.

✓ Where individualized dosing (dose and/or duration) is based on a biomarker responsive to treatment.
1. Some Terminology

Targeted therapeutics means Finding subgroups
2. Clarifying Goals
2. Clarifying Goals

Targeted therapeutics means finding subgroups. What’s real versus what’s spurious?
2. Clarifying Goals

Do you want to find a subgroup or not?

**YES** – Heterogeneity is my friend!
- I want to find a targeted therapeutic!
Clarifying Goals

Do you want to find a subgroup or not?

**YES** – Heterogeneity is my friend!
  - I want to find a targeted therapeutic!

**NO** – Heterogeneity is my enemy!
  - I want the treatment effect to be homogeneous across subgroups.
3. Disciplined Subgroup Search (DSS)

3. Disciplined Subgroup Search

Definition

Disciplined Subgroup Search (DSS) is

- a systematic approach to subgroup identification having
- several characteristics that are quite distinct
  - from those of a traditional subgroup analysis
  - or post hoc exploratory analysis of subgroups.
DSS characteristics

1. **Prespecification**: the algorithm/methodology to be used for identifying subgroups, the list of biomarkers that form the covariate space to be searched, complexity of subgroup definitions (i.e., how many covariates are allowed to define the subgroup), as well as any other options/decisions that can be made in the analysis process.

   - In short, this is no different than prespecification of any important analysis in a Phase 3 trial that adheres to the ICH-E9 Guideline.
3. Disciplined Subgroup Search

DSS characteristics

2. Adjusting for multiplicity: how statistical significance (i.e., $p$-values) of a subgroup finding will be adjusted for multiplicity. [Also consider Bayesian approaches.]

3. Bias correction: how estimates of treatment effect are corrected for bias due to the selection bias associated with searching multiple subgroups.
3. Disciplined Subgroup Search

DSS characteristics

4. Biomarker effects: allows for separating prognostic biomarker effects from predictive biomarker effects.

5. Interactions: allows for multiple biomarkers to be included in the definition of a subgroup.

6. Partition: allows for identification of a cut-off value for a continuous biomarker that separates smaller treatment effects from larger treatment effects.

[Features added to Ilya’s table?]
4. Always Do Subgroup Identification (DSS)
Subgroup Identification

“Always do subgroup analysis, but never believe them.”

Attributed to Sir Richard Peto
Professor of Medical Statistics and Epidemiology
University of Oxford, England
Subgroup Identification

Do You Want To Find a Subgroup or Not?

NO

YES
Subgroup Identification

Do You Want To Find a Subgroup or Not?

NO

“Routine” baseline factors

Playing “defense”

Avoid trying to explain away (unusual?) findings

Use DSS !!
Subgroup Identification

Do You Want To Find a Subgroup or Not?

NO

“Routine” baseline factors
Playing “defense”
Avoid trying to explain away (unusual?) findings
Use DSS !!
“The figure ... illustrates principal results for a wide variety of subgroup comparisons, including US vs. non-US populations (the latter of which was not pre-specified). The combined endpoints of all-cause mortality plus all-cause hospitalization and of mortality plus heart failure hospitalization showed consistent effects in the overall study population and the subgroups, including women and the US population. However, in the US subgroup and women, overall mortality and cardiovascular mortality appeared less affected. Analyses of female and US patients were carried out because they each represented about 25% of the overall population. Nonetheless, subgroup analyses can be difficult to interpret, and it is not known whether these represent true differences or chance effects.”

From US Label for Toprol-XL
Subgroup Identification

Ticagrelor Example

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Patients</th>
<th># of Events</th>
<th>HR (95% CI)</th>
<th>Interaction p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographic Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia / Australia</td>
<td>1714</td>
<td>95</td>
<td>116</td>
<td>0.80 (0.61, 1.04)</td>
</tr>
<tr>
<td>Cent / Sth America</td>
<td>1237</td>
<td>91</td>
<td>104</td>
<td>0.86 (0.65, 1.13)</td>
</tr>
<tr>
<td>Euro / Md E / Afr</td>
<td>13859</td>
<td>576</td>
<td>712</td>
<td>0.80 (0.72, 0.90)</td>
</tr>
<tr>
<td>North America</td>
<td>1814</td>
<td>102</td>
<td>82</td>
<td>1.25 (0.93, 1.67)</td>
</tr>
</tbody>
</table>

Source: Sponsor presentation at CV and Renal Drugs Ad Comm Meeting July, 2010 CC-30
Always do subgroup identification !!!

What if DSS had been formally done?

- Often subgroups defined by baseline factors are described in the protocol (e.g. gender, race, baseline severity, etiology, etc.?)
- What if the subgroup identification search methodology was pre-specified?
- What if adjusted p-values and effect estimates were calculated?
- Would these surprising findings not be so surprising anymore?
Subgroup Identification

Do You Want To Find a Subgroup or Not?

YES

Many baseline factors, biomarkers, med Hx, etc.

Playing “offense”

Seeking to explain findings

Use DSS !!
Subgroup Identification

Do You Want To Find a Subgroup or Not?

YES

Many baseline factors, biomarkers, med Hx, etc.

Playing “offense”

Seeking to explain findings

Use DSS !!

Ramucirumab vs Placebo in HCC (REACH)
Lancet Onc, 2015; 16, 859-870
Subgroup Identification

Do You Want To Find a Subgroup or Not?

Many baseline factors, biomarkers, med Hx, etc.

Playing “offense”

Seeking to explain findings

Use DSS !!
Do You Want To Find a Subgroup or Not?

YES

Many baseline factors, biomarkers, med Hx, etc.

Playing “offense”

Seeking to explain findings

Use DSS!!

Ramucirumab vs Placebo in HCC (REACH)
Lancet Onc, 2015; 16, 859-870
Subgroup Identification

Do You Want To Find a Subgroup or Not?

Ramucirumab vs Placebo in HCC with elevated AFP (REACH-2)
Lancet Onc, 2019; 20, 282-296

Many baseline factors, biomarkers, med Hx, etc.

Playing “offense”

Seeking to explain findings

Use DSS !!
Always do subgroup identification !!!

What if DSS had been formally done in REACH?

- What if the AFP subgroup was pre-specified along with other subgroups?
- What if the subgroup identification search methodology was pre-specified?
- What if adjusted p-values and effect estimates were calculated?
- What if they were still significant and meaningful?
Always do subgroup identification !!!

What if DSS had been formally done in REACH?
- What if the AFP subgroup was pre-specified along with other subgroups?
- What if the subgroup identification search methodology was pre-specified?
- What if *adjusted* p-values and effect estimates were calculated?
- What if they were still significant and meaningful?

But ... DSS is hard !!!!!
The $1,000,000,000 [read billion] question...

Could ramucirumab have been approved in the targeted subgroup based on REACH in 2015 using DSS instead of

- Spending 3 years, and
- Many, many millions of dollars, and
- Tens of thousands of patients not having access to an effective medication?
The $1,000,000,000 [read billion] question ...

Could ramucirumab have been approved in the targeted subgroup based on REACH in 2015 instead of

- Spending 3 years, and
- Many, many millions of dollars, and
- Tens of thousands of patients not having access to an effective medication?

Is DSS a billion dollars hard?!?!?!?
“Always do subgroup *identification using DSS* so the results are more believable.”

Steve Ruberg

Your Run-of-the-Mill Statistician
5. One More Idea
Suppose there are 100 potential predictive biomarkers that could be important for a new treatment.
Suppose there are 100 potential predictive biomarkers that could be important for a new treatment.

Observed p-value = 0.0001 for one biomarker

Bonferroni adjusted p-value $\leq 100 \times 0.0001 = 0.01$
Suppose there are 100 potential predictive biomarkers that could be important for a new treatment.

Observed p-value = 0.0001 for one biomarker

- Bonferroni adjusted p-value $\leq 100 \times 0.0001 = 0.01$

EUREKA! We have discovered a novel biomarker-defined subgroup.
ARE YOU SURE?
**ARE YOU SURE?**

Suppose further

\[
\text{pr(success ... finding a biomarker)} = \text{pr(at least one } H_0 \text{ is false)} = 0.20
\]

Prior on \( H_0 \) is true (none are predictive) = 0.80
ARE YOU SURE?

Suppose further

\[
\text{pr}(\text{success ... finding a biomarker}) = \text{pr}(\text{at least one } H_0 \text{ is false}) = 0.20
\]

Prior on \( H_0 \) is true (none are predictive) = 0.80

Uniform prior per biomarker = \( \frac{0.20}{100} = 0.002 \)
One More Idea

Let $p_0 = \text{prior probability that } H_0 \text{ is false (e.g. the biomarker is predictive)}$

Let $p = \text{observed p-value for test statistics for } H_0$

Bayes factor* $\left[ -e \times p \times \ln(p) \right]^{-1}$ can be used to give an upper bound on the posterior probability that $H_0$ is false.

One More Idea

Let $p_0 = \text{prior probability that } H_0 \text{ is false (e.g. the biomarker is predictive)}$

Let $p = \text{observed p-value for test statistics for } H_0$

Bayes factor* $[-e \times p \times \ln(p)]^{-1}$ can be used to give an upper bound on the posterior probability that $H_0$ is false

Posterior probability** for $H_0$ being false ($p_1$) is (upper bound)

$$p_1 \leq \{1 + [(1-p_0)/p_0] \times [-e \times p \times \ln(p)] \}^{-1}$$


**If $p < 1/e = .368$
ARE YOU SURE?

Suppose further

\[ \text{pr}(\text{success ... finding a biomarker}) = \text{pr}(\text{at least one } H_0 \text{ is false}) = 0.20 \]

Prior on \( H_0 \text{ is true (none are predictive)} = 0.80 \)

Uniform prior per biomarker = \( \frac{0.20}{100} = 0.002 \)

Bayesian posterior \( \text{pr}(H_0 \text{ is false}) \leq 0.44. \)
Subgroup Identification

“Always use Bayesian thinking when doing subgroup identification so you can quantify how believable the results are.”

Steve Ruberg
Your Run-of-the-Mill Bayesian Statistician
Conclusion
Subgroup identification is the HOLY GRAIL.

Not surprisingly, that makes it **the hardest problem there is**.

- Size of the subgroup (10, 20, ... 50%)
- Differential effect in the subgroup (5, 15, 25%)
- No. of biomarkers (3, 10, 100, 1000, ... $10^6$)
- No. of biomarkers defining the subgroup (1, 2, 3, ...)
- Nature of biomarker effect (step or smooth curve)
- Size of the clinical trial (100, 250, 500, 1000/group)
Conclusion

ALWAYS do subgroup identification ...

To find or not to find ...

Disciplined Subgroup Search

(replication and biological plausibility are very important)
Conclusion

ALWAYS do subgroup identification ...

To find or not to find ...

Disciplined Subgroup Search

When I see a significant subgroup finding,
I always ask ...

“I wonder what their prior was?”
THANK YOU

$X > a$

$X < a$