Worked example of visualization tools when estimating effects in target populations

Karolinska Institutet
Tuesday, November 8, 2022

General breakdown

- Variable importance to transport figure
  - Breakout group session #1-7 minutes
- Diagnosing potential problems with weighted analyses
  - Breakout group session #2-10 minutes
- Performing and interpreting results
- General discussion

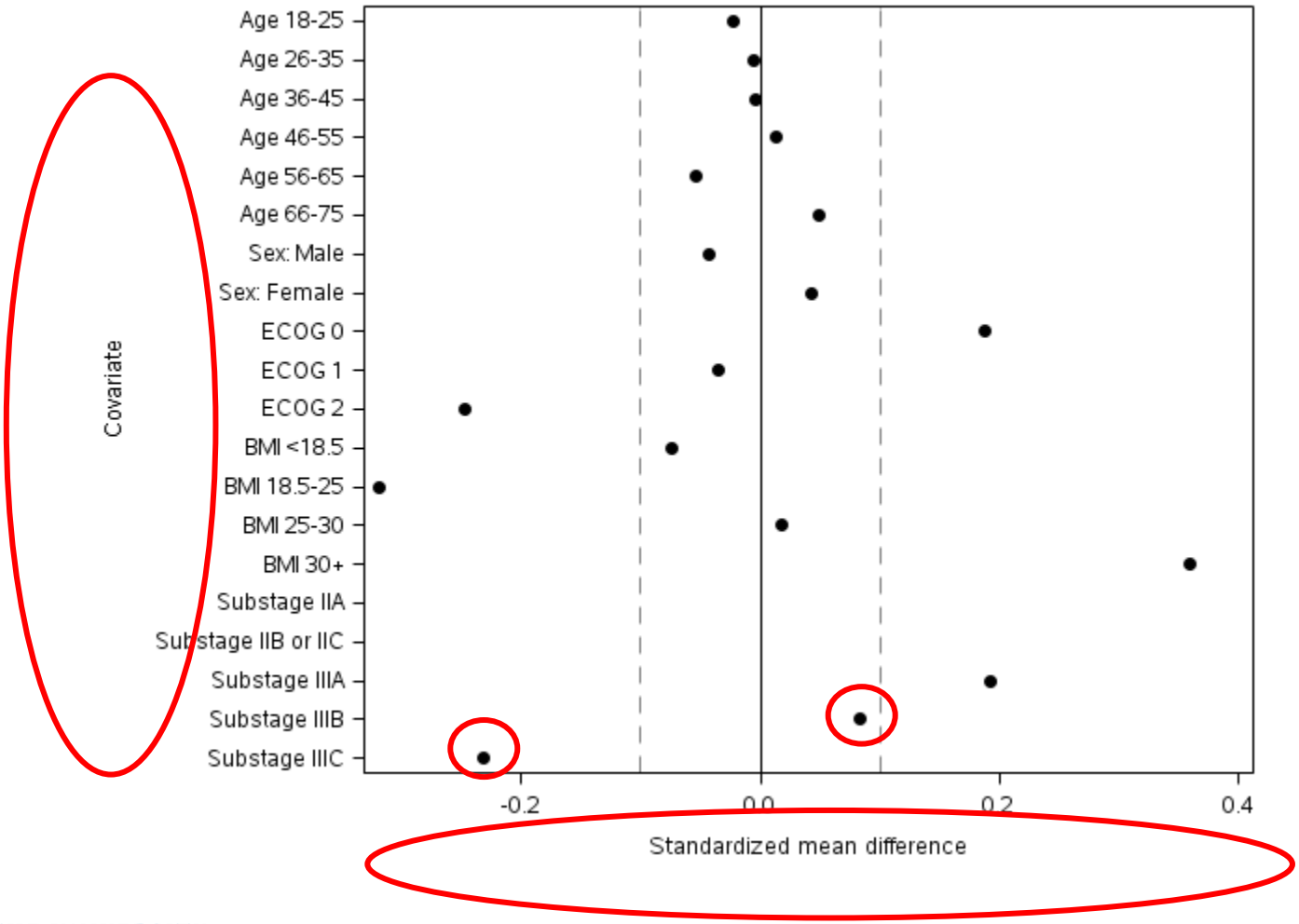
Variable selection figures

1. **Task:** Identify a set of variables that may be conditionally associated with the outcome based on existing knowledge
   Approach: Review of literature, clinical input, or use of directed acyclic graphs/selection diagrams
2. **Task:** Conduct a high-level screening for potential effect measure modifiers within trial and target data sets
   Approach: Variable Importance for Treatment Transport (VITT) plot **(Figure 1)**
3. **Task:** Implement a more flexible trial sampling model including identified effect measure modifiers
   Approach: Allow model to incorporate interactions and other non-linearities
4. **Task:** Evaluate external positivity assumption
   Approach: Evaluate distribution of trial sampling probabilities and overlap between trial and target **(Figure 2)**
5. **Task:** Assess balance of effect measure modifiers after applying inverse odds of sampling weights
   Approach: Assess balance of effect measure modifiers in the trial and target before and after weighting **(Figure 3)**
6. **Task:** Identify potentially influential observations with extreme weights
   Approach: Evaluate distribution of inverse odds of sampling weights in the weighted trial population **(Figure 4)**
7. **Task:** De-brief on steps 1-6 above.
   Approach: Evaluate the plausibility and impact of assumptions and consider implications of the output.

Variable selection/importance and external validity

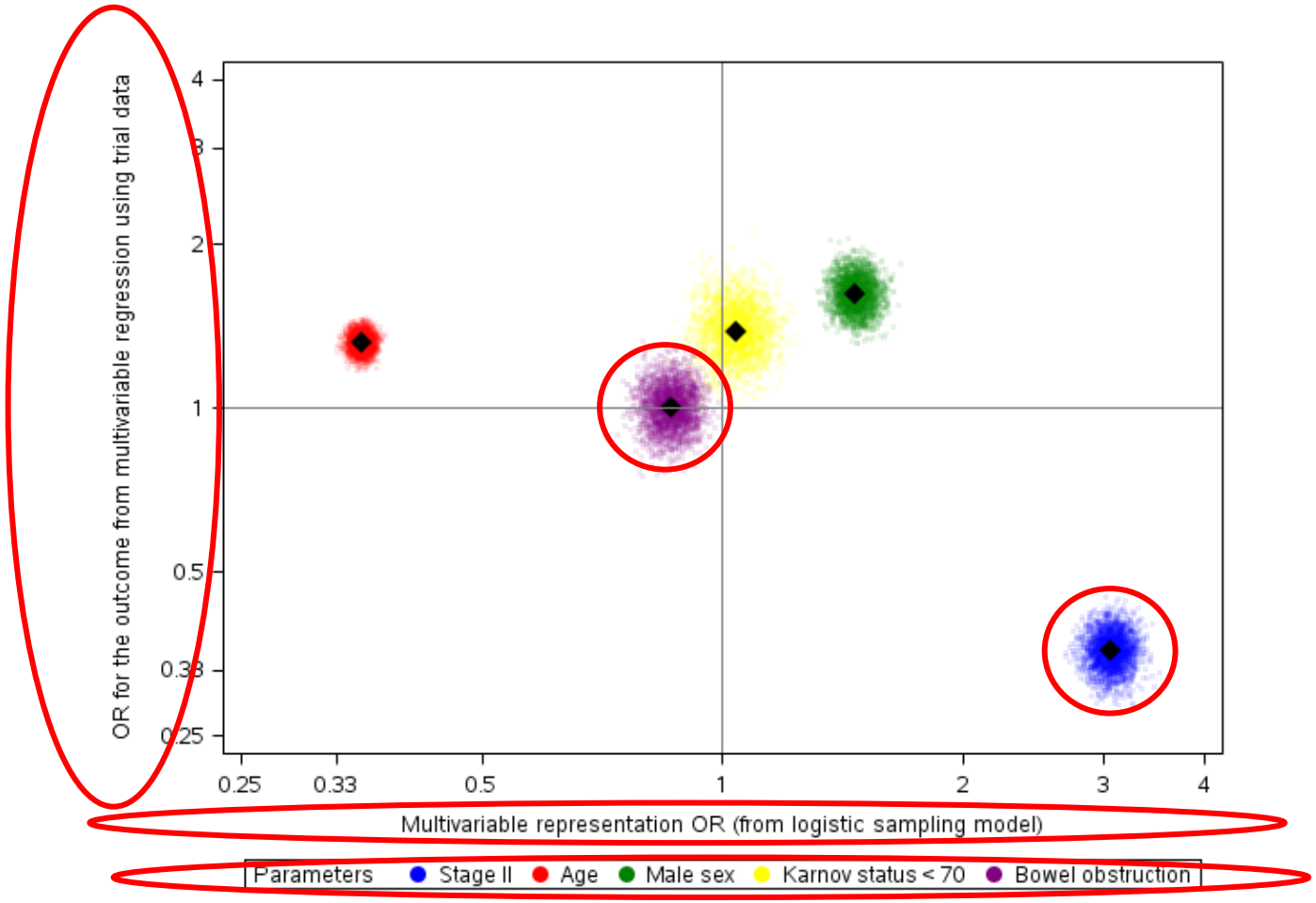
- Remember that whichever method we use, we need to select variables for an adjustment set
- There is ongoing work on using directed acyclic graphs and selection diagrams for this purpose
- Due to their non-parametric nature, these tools say very little about the relative importance of each variable
- A new plot (VITT plot) of multivariable relationships can give some insight here (building on an existing plot, the LOVE plot)
  - These plots are **method-agnostic**

Love plots of standardized mean differences

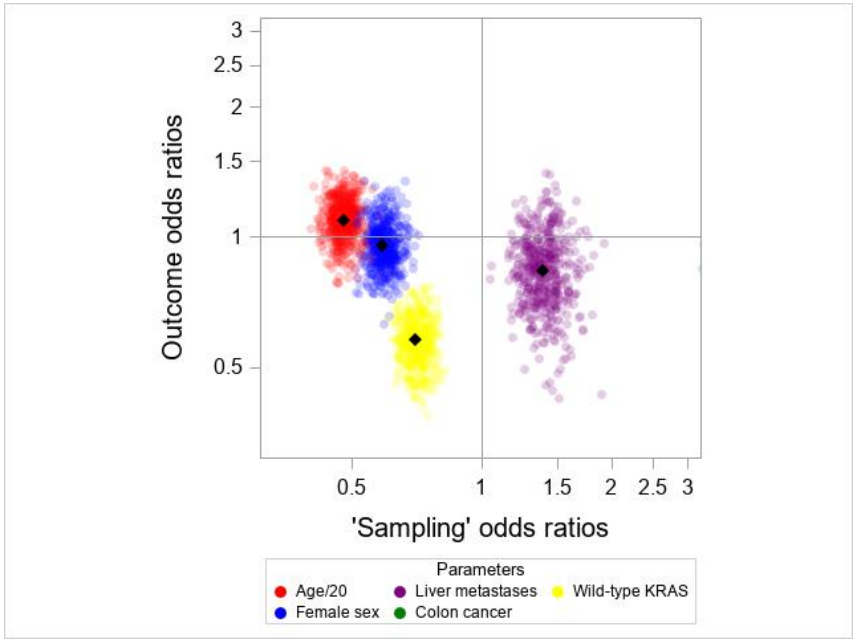


Code to reference: Love_plot_code

VITT plots of "sampling" and outcome ORs



Discussion time: VITT plot



Code to create this plot: VITT_plot_code

Questions:

1) What variables seem most important for inclusion in the adjustment set?
2) What variables seem the least important?
3) Are there any variables you would consider dropping completely?

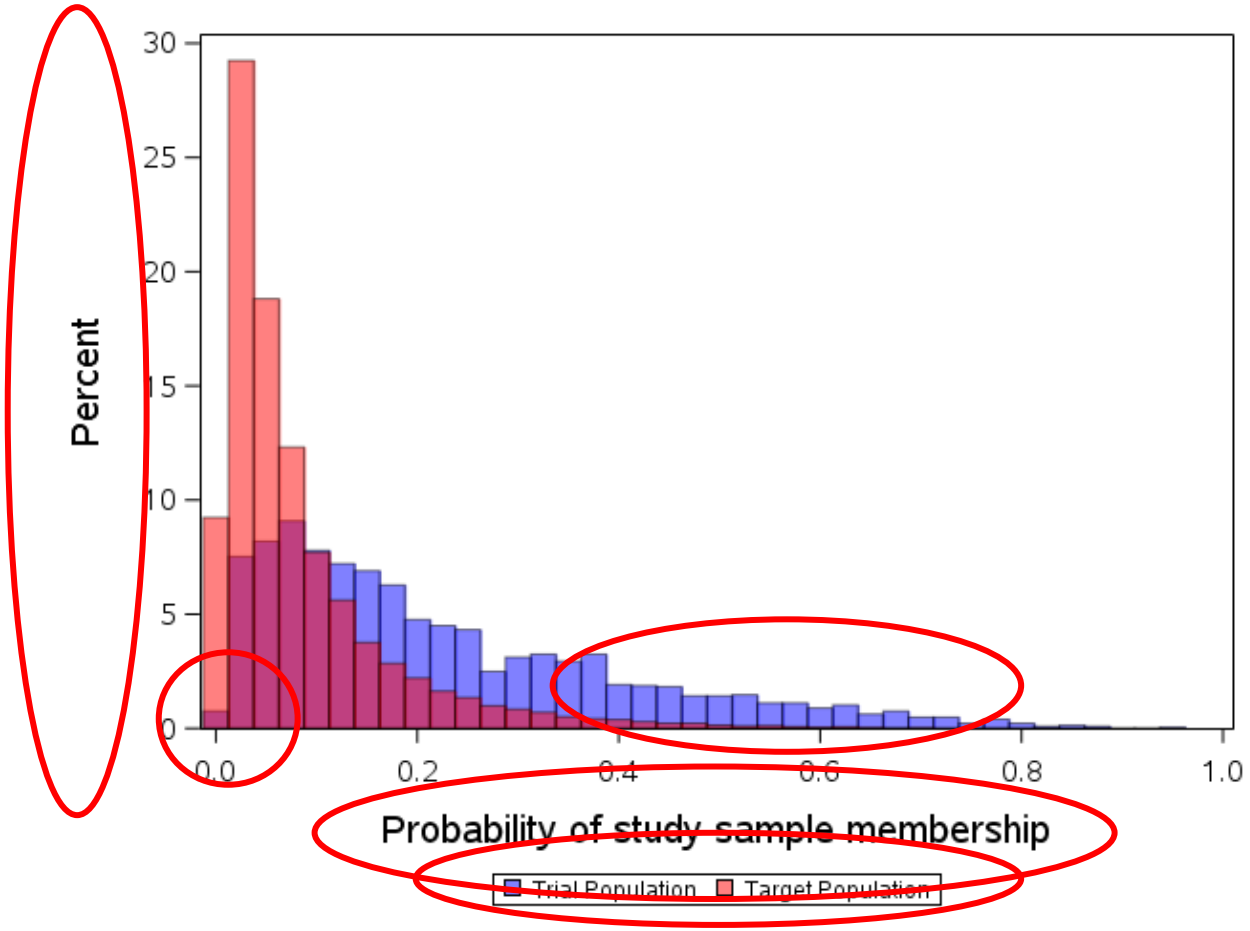
Diagnosing potential problems

Solving the external validity problem with weights

- Once you have identified the key variables, it is necessary to choose an analytic method
- Here we'll be focusing on weights and potential issues that can arise (especially the lack of transport positivity)
- When using these weights, some preliminary visualizations can help diagnose potential problems
  - Histograms and density plots of estimated probabilities
  - Pre- and post-weighting Love plots
  - "Skyscraper" plots of the weights to be used

1. **Task:** Identify a set of variables that may be conditionally associated with the outcome based on existing knowledge
   Approach: Review of literature, clinical input, or use of directed acyclic graphs/selection diagrams
2. **Task:** Conduct a high-level screening for potential effect measure modifiers within trial and target data sets
   Approach: Variable Importance for Treatment Transport (VITT) plot **(Figure 1)**
3. **Task:** Implement a more flexible trial sampling model including identified effect measure modifiers
   Approach: Allow model to incorporate interactions and other non-linearities
4. **Task:** Evaluate external positivity assumption
   Approach: Evaluate distribution of trial sampling probabilities and overlap between trial and target **(Figure 2)**
5. **Task:** Assess balance of effect measure modifiers after applying inverse odds of sampling weights
   Approach: Assess balance of effect measure modifiers in the trial and target before and after weighting **(Figure 3)**
6. **Task:** Identify potentially influential observations with extreme weights
   Approach: Evaluate distribution of inverse odds of sampling weights in the weighted trial population **(Figure 4)**
7. **Task:** De-brief on steps 1-6 above.
   Approach: Evaluate the plausibility and impact of assumptions and consider implications of the output.

Histograms of “sampling” probabilities

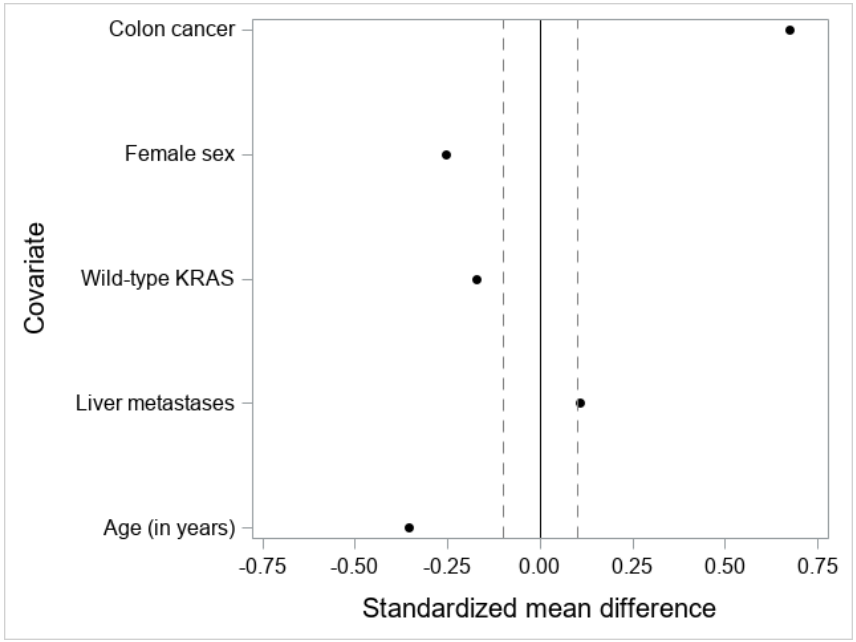


Code to reference: Density_plot_code

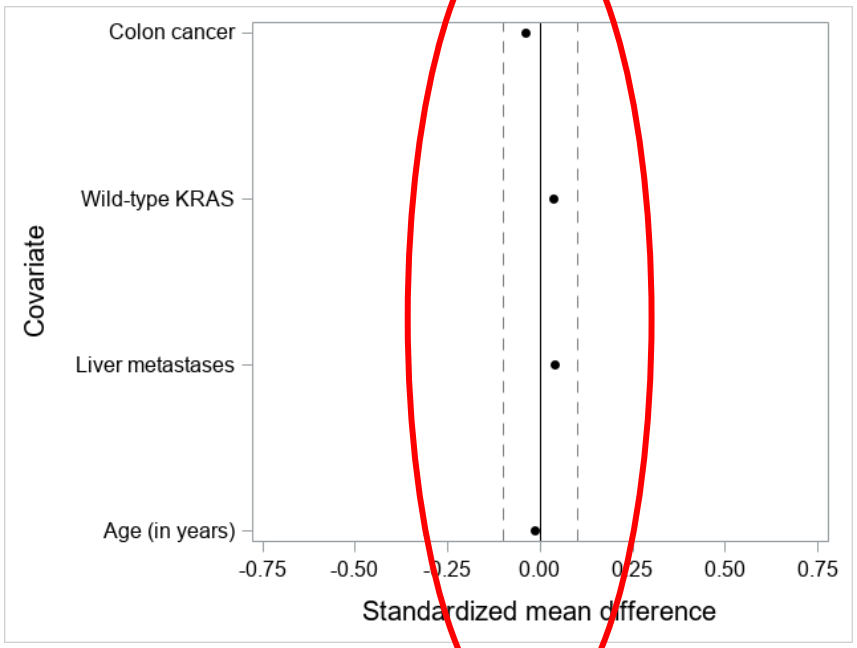
1. **Task:** Identify a set of variables that may be conditionally associated with the outcome based on existing knowledge
   Approach: Review of literature, clinical input, or use of directed acyclic graphs/selection diagrams
2. **Task:** Conduct a high-level screening for potential effect measure modifiers within trial and target data sets
   Approach: Variable Importance for Treatment Transport (VITT) plot **(Figure 1)**
3. **Task:** Implement a more flexible trial sampling model including identified effect measure modifiers
   Approach: Allow model to incorporate interactions and other non-linearities
4. **Task:** Evaluate external positivity assumption
   Approach: Evaluate distribution of trial sampling probabilities and overlap between trial and target **(Figure 2)**
5. **Task:** Assess balance of effect measure modifiers after applying inverse odds of sampling weights
   Approach: Assess balance of effect measure modifiers in the trial and target before and after weighting **(Figure 3)**
6. **Task:** Identify potentially influential observations with extreme weights
   Approach: Evaluate distribution of inverse odds of sampling weights in the weighted trial population **(Figure 4)**
7. **Task:** De-brief on steps 1-6 above.
   Approach: Evaluate the plausibility and impact of assumptions and consider implications of the output.

Love plots to examine balance

Before inverse odds weights

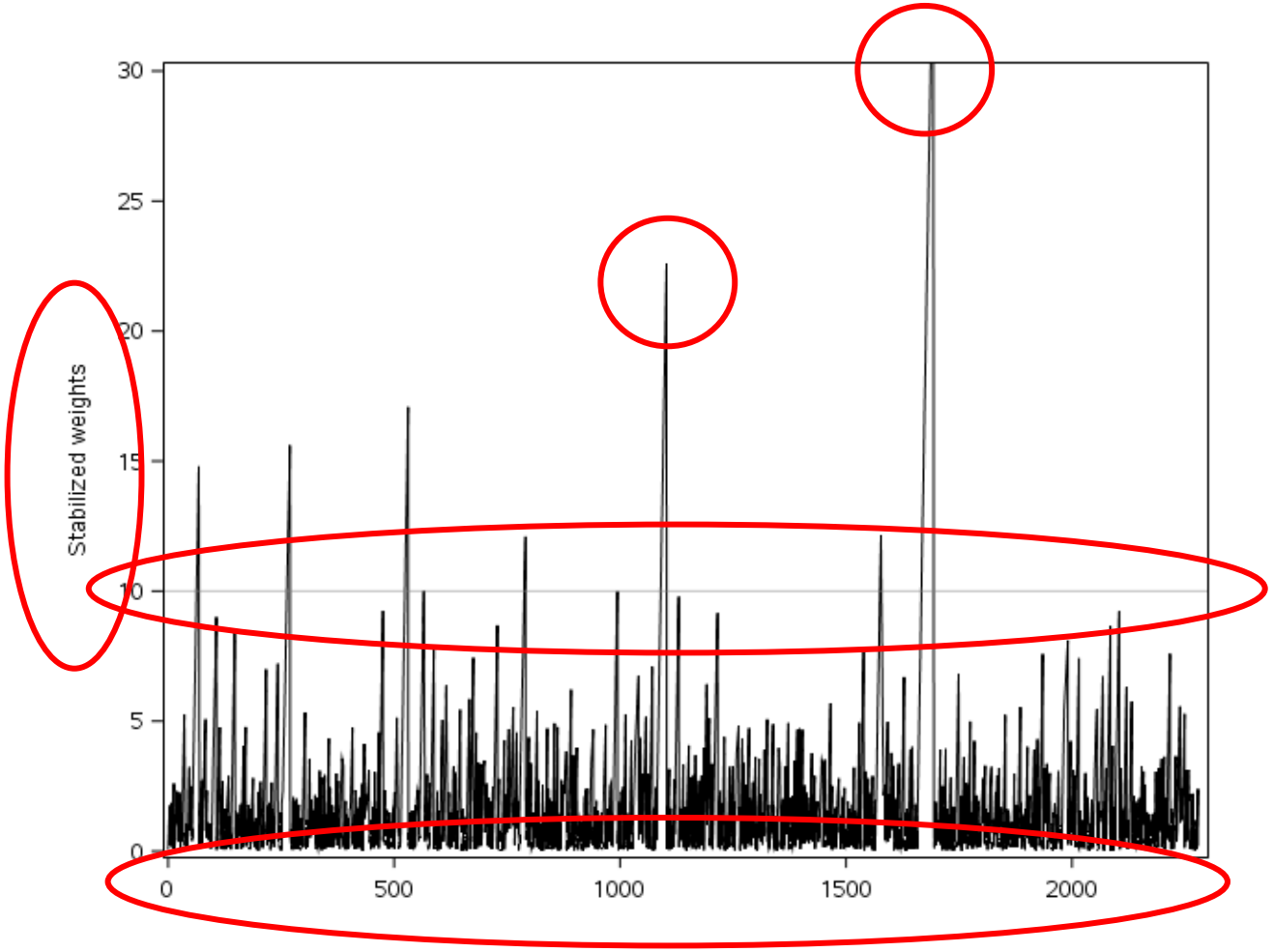


After inverse odds weights



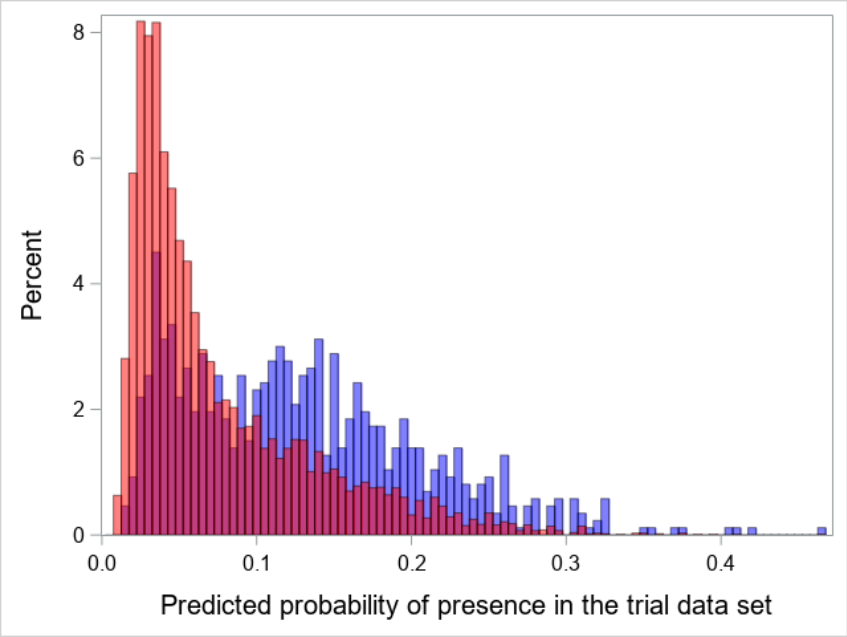
1. **Task:** Identify a set of variables that may be conditionally associated with the outcome based on existing knowledge
   Approach: Review of literature, clinical input, or use of directed acyclic graphs/selection diagrams
2. **Task:** Conduct a high-level screening for potential effect measure modifiers within trial and target data sets
   Approach: Variable Importance for Treatment Transport (VITT) plot **(Figure 1)**
3. **Task:** Implement a more flexible trial sampling model including identified effect measure modifiers
   Approach: Allow model to incorporate interactions and other non-linearities
4. **Task:** Evaluate external positivity assumption
   Approach: Evaluate distribution of trial sampling probabilities and overlap between trial and target **(Figure 2)**
5. **Task:** Assess balance of effect measure modifiers after applying inverse odds of sampling weights
   Approach: Assess balance of effect measure modifiers in the trial and target before and after weighting **(Figure 3)**
6. **Task:** Identify potentially influential observations with extreme weights
   Approach: Evaluate distribution of inverse odds of sampling weights in the weighted trial population **(Figure 4)**
7. **Task:** De-brief on steps 1-6 above.
   Approach: Evaluate the plausibility and impact of assumptions and consider implications of the output.

“Skyscraper” plots to identify high weights

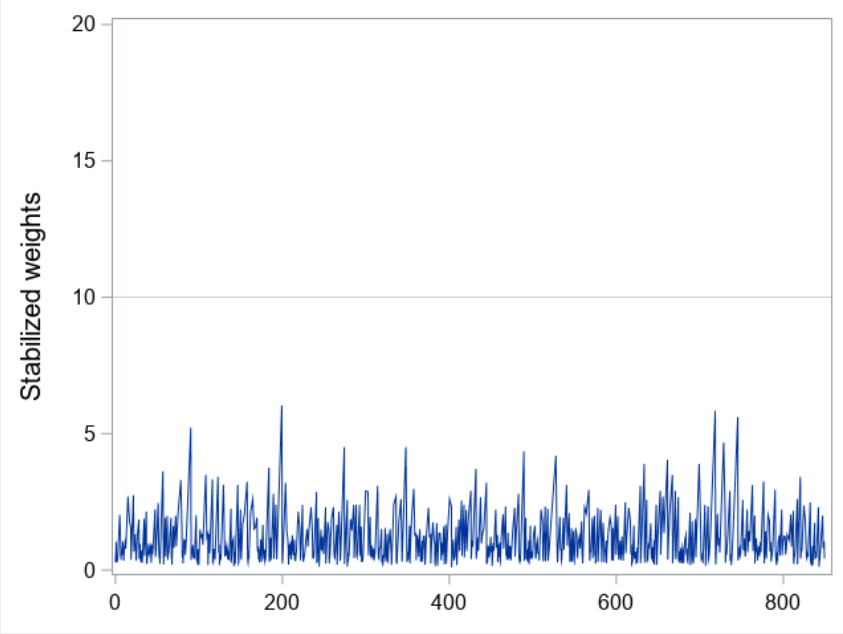


Code to reference: Skyscraper_plot_code

Discussion time: probability and skyscraper plots



Code to create this plot: Density_plot_code



Code to create this plot: Skyscraper_plot_code

Questions:

1) Do you see any potential problems with using this model based on the density plot?
2) What about the skyscraper plot? Do you see any potentially problematic weights?
3) Purely based on these plots, would you be comfortable proceeding with the analysis?

Analyzing data

1. **Task:** Identify a set of variables that may be conditionally associated with the outcome based on existing knowledge
   Approach: Review of literature, clinical input, or use of directed acyclic graphs/selection diagrams
2. **Task:** Conduct a high-level screening for potential effect measure modifiers within trial and target data sets
   Approach: Variable Importance for Treatment Transport (VITT) plot **(Figure 1)**
3. **Task:** Implement a more flexible trial sampling model including identified effect measure modifiers
   Approach: Allow model to incorporate interactions and other non-linearities
4. **Task:** Evaluate external positivity assumption
   Approach: Evaluate distribution of trial sampling probabilities and overlap between trial and target **(Figure 2)**
5. **Task:** Assess balance of effect measure modifiers after applying inverse odds of sampling weights
   Approach: Assess balance of effect measure modifiers in the trial and target before and after weighting **(Figure 3)**
6. **Task:** Identify potentially influential observations with extreme weights
   Approach: Evaluate distribution of inverse odds of sampling weights in the weighted trial population **(Figure 4)**
7. **Task:** De-brief on steps 1-6 above.
   Approach: Evaluate the plausibility and impact of assumptions and consider implications of the output.

Weighted analytics

- Having completed our diagnostics, how do we actually estimate a treatment effect?
- PROC PHREG conveniently allows use of a “WEIGHT” statement, as does many R packages for estimating hazard ratios
- Notably, traditional tools for estimating variance can ignore variability from sampling your target population
- Instead, we have bootstrapped, re-drawing from **both** the trial and target each iteration
  - And balancing each treatment arm separately

Final results

Target population	Progression-free survival hazard ratio	95% confidence limits	Confidence limit ratio
PRIME trial	0.97	0.84, 1.11	1.33
Base target population	0.86	0.74, 0.99	1.34
Even higher KRAS target population	**0.79**	0.67, 0.93	1.40

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Questions?

mawc@live.unc.edu