numerus
Discover Statistical Power

Comparing ophthalmology treatments via the integration of IPD and aggregate-level data:

A new MAIC approach

Jonathan Alsop

Director

Numerus

PSI Annual Conference, 5th June 2018

EXPERT REVIEW OF PHARMACOECONOMICS & OUTCOMES RESEARCH, 2016
http://dx.doi.org/10.1586/14737167.2016.1165609

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

Review and comparison of methodologies for indirect comparison of clinical trial results: an illustration with ranibizumab and aflibercept

Stephane A. Regnier[a], Jonathan Alsop[b], Jonathan Wright[b], Richard Nixon[a], Harry Staines[a] and Franck Fajnkuchen[c,d]

[a]Novartis Pharma AG, Basel, Switzerland; [b]Numerus Ltd, Wokingham, UK; [c]Centre Ophtalmologique d'Imagerie et de Laser, Paris, France; [d]Hôpital Avicenne, Assistance Publique Hôpitaux de Paris, Bobigny, France

ABSTRACT

Aim: To review and compare methods for indirect comparison of aflibercept and ranibizumab in patients with diabetic macular edema.

Methods: Post-stratification, inverse probability weighting based on simulated data, weight optimization, and regression model techniques were used to compare pooled individual patient-level data from the RESTORE and RESPOND (ranibizumab 0.5 mg as needed after 3 initial monthly doses) studies with summary-level data from the VIVID and VISTA (aflibercept 2.0 mg every 8 weeks after 5 initial monthly doses, 2q8) studies. The impact of adjusting for up to two baseline characteristics was assessed.

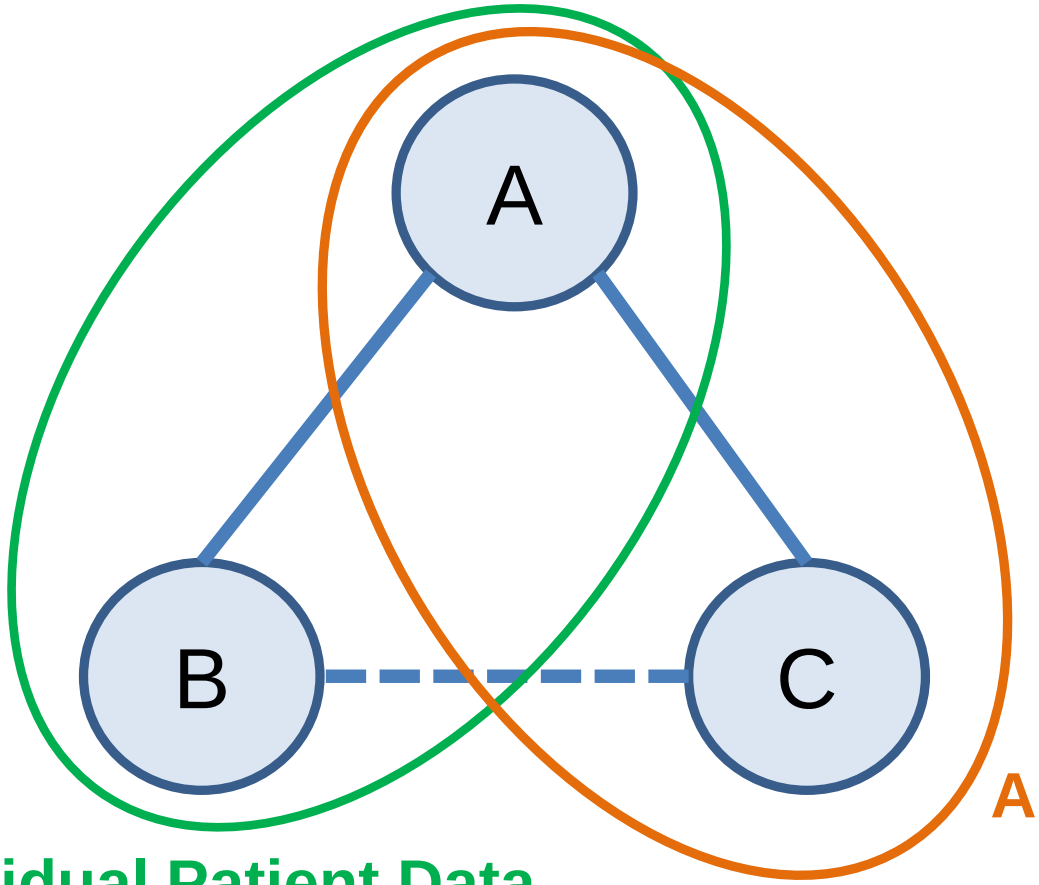
Results: All methods provided similar results. After adjustment for baseline best-corrected visual acuity and central retinal thickness, no statistically significant difference in average gain in baseline best-corrected visual acuity from baseline at month 12 was found between ranibizumab 0.5 mg and aflibercept 2q8.

Conclusions: Weight optimization and regression methods are useful options to adjust for more than one baseline characteristic.

ARTICLE HISTORY
Received 23 December 2015
Accepted 10 March 2016
Published online
25 March 2016

KEYWORDS
Confounders; indirect treatment comparison; inverse probability weighting; regression; stratified sampling

Indirect treatment comparison



Individual Patient Data (IPD)

Aggregate-level Data (AGR)

A ‘baseline’ problem….

numerus
Discover Statistical Power

≠

Trial #1

Individual Patient Data

Trial #2

Aggregate Data

What is MAIC?

numerus
Discover Statistical Power

Matching **A**djusted **I**ndirect **C**omparison

Re-weighting of IPD

Weighted IPD baseline characteristics =
AGR baseline characteristics

Weighted analysis of IPD outcomes

c.f. with AGR outcome data

Optimization Challenge

numerus
Discover Statistical Power

Suppose i^{th} patient has a_i (baseline measure)

Primary: find weights w_i

$\sum_i w_i a_i = \bar{a}_{(2)}$ Target mean in AGR

$\sum_i w_i = n_1$ IPD study size

$w_i \geq 0$

Secondary:

$w_i \leq w_{max}$ Restrict weights (avoid possible high leverage)

$\frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2}$ Maximise Effective Sample Size (ESS)

Methods

- Signorovitch et al. (2010)
- Post-stratification (Di Lorenzo et al., 2011)
- Entropy balancing (Hainmueller, 2012)
- MAICn (Han, 2014)
- **Polynomial weighting (Regnier et al. 2015)**

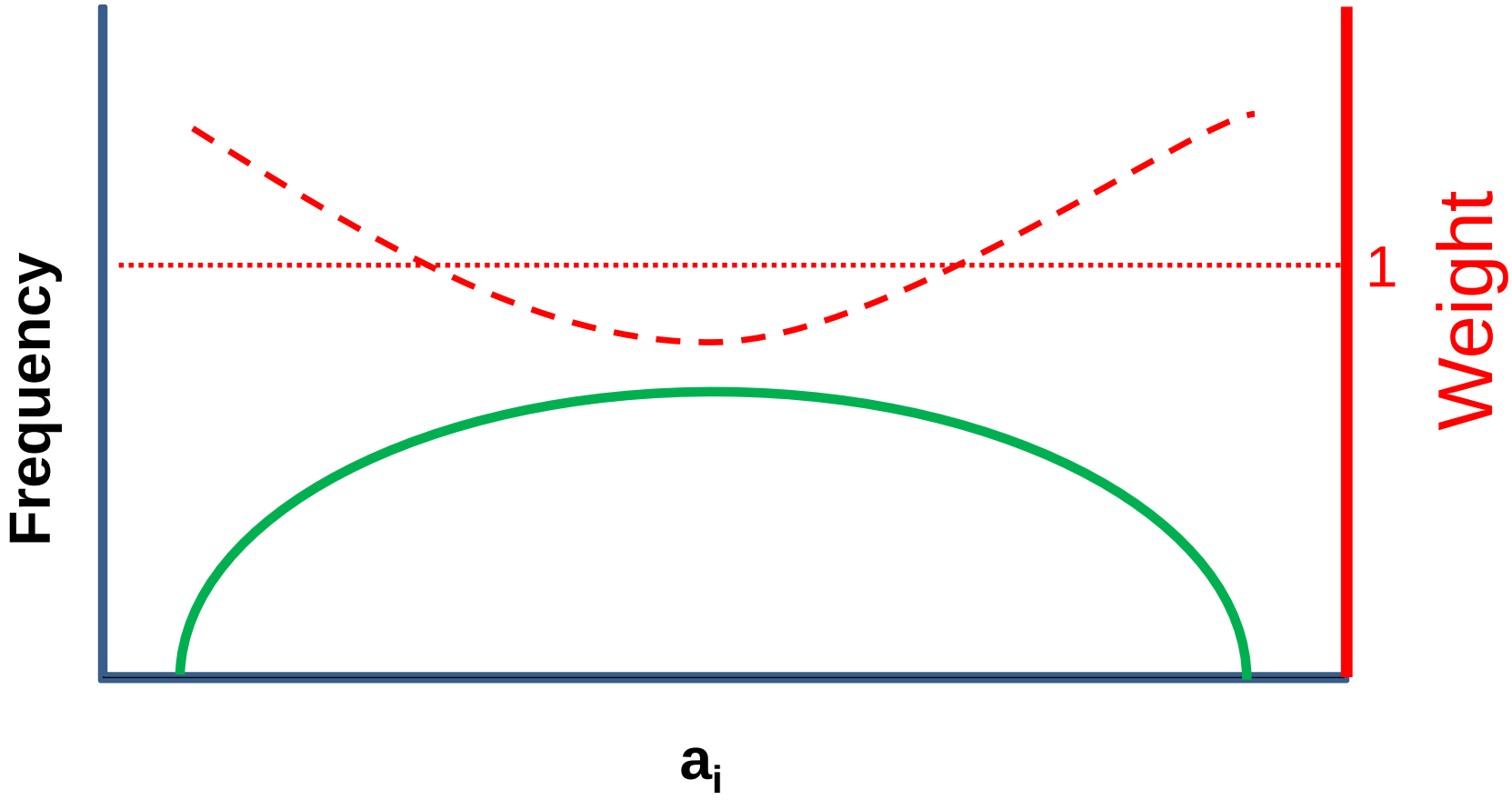
Borrowing from historic controls

Survey sampling

(e.g. cell weighting, raking)

Polynomial weighting

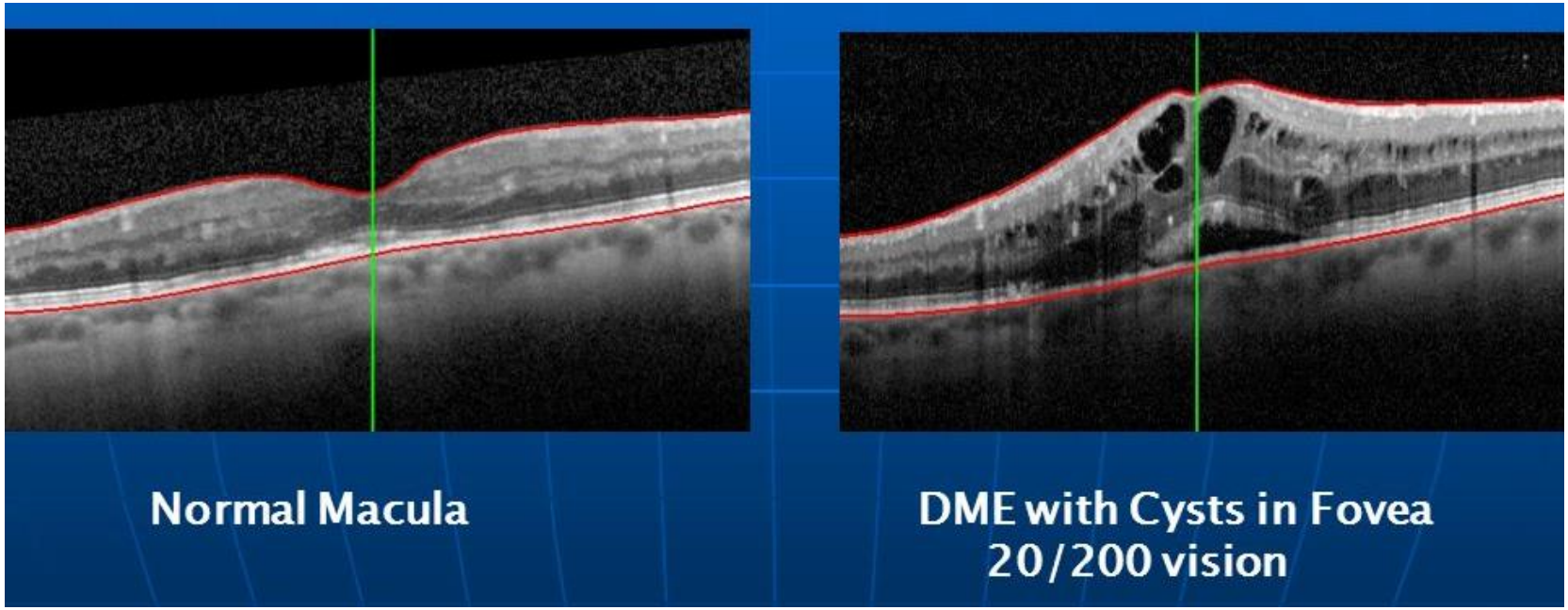
Polynomial weighting



Example

numerus
Discover Statistical Power

Diabetic Macular Edema (DME)

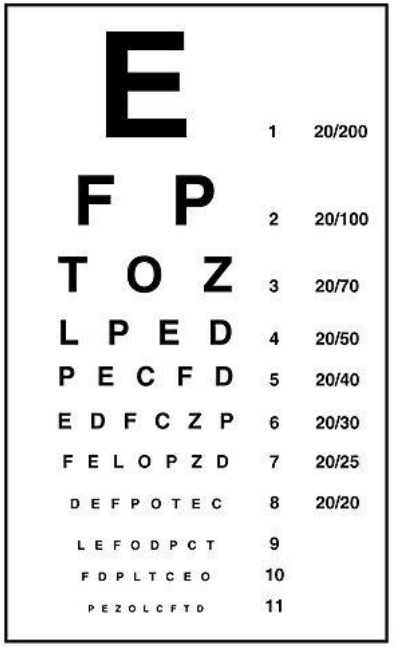


Comparison of two Anti-VEGF treatments

Studies

numerus
Discover Statistical Power

Characteristic	Ranibizumab	Aflibercept
Data type	IPD	AGR
Sponsor	Novartis	Regeneron
Studies	RESTORE RESPOND	VIVID VISTA
Study type	Blinded RCT	
Control arm	Laser photocoagulation therapy	
Endpoint	Δ from Baseline to M12/LOCF in visual acuity (letters)	



Treatment confounders

numerus
Discover Statistical Power

Baseline Characteristic	Aflibercept arm (N=286)	Ranibizumab arm (N=190)	Difference (95% CI)
BCVA (letters)	59.1 (11.0)	64.1 (10.3)	-5.0 (-7.0,-3.0)
CRT (µm)	497 (151)	435 (126)	62 (36,88)

What would have been effect of Ranibizumab in the VIVID/VISTA population?

Polynomial weighting

numerus
Discover Statistical Power

j=baseline parameter, i=patient

1) a_{ij} standardised to lie on [0,1]: $$a^*_{ij} = \frac{a_{ij} - \min(a_{ij})}{\max(a_{ij}) - \min(a_{ij})}$$

2) $$w^*_{ij} = \left|\beta_{2j} + \beta_{3j}\left(a^*_{ij} - \beta_{1j}\right) + \beta_{4j}\left(a^*_{ij} - \beta_{1j}\right)^2 + \beta_{5j}\left(a^*_{ij} - \beta_{1j}\right)^3 + \beta_{6j}\left(a^*_{ij} - \beta_{1j}\right)^4\right|$$

3) $$w_i = \frac{\sum_j w^*_{ij}}{\sum_{ij} w^*_{ij}}$$

4) Calculate weighted mean & SD of a_{ij} using $w_i \rightarrow Mw_j$, SDw_j

5) Minimise $L = \sum_j \frac{\left[\left(Mw_j - Mt_j\right)^2 + \left(SDw_j - SDt_j\right)^2\right]}{SDt_j}$ w.r.t. β_{ij}

Mt_j , SDt_j = target mean & SD (AGR)

Baseline Matching

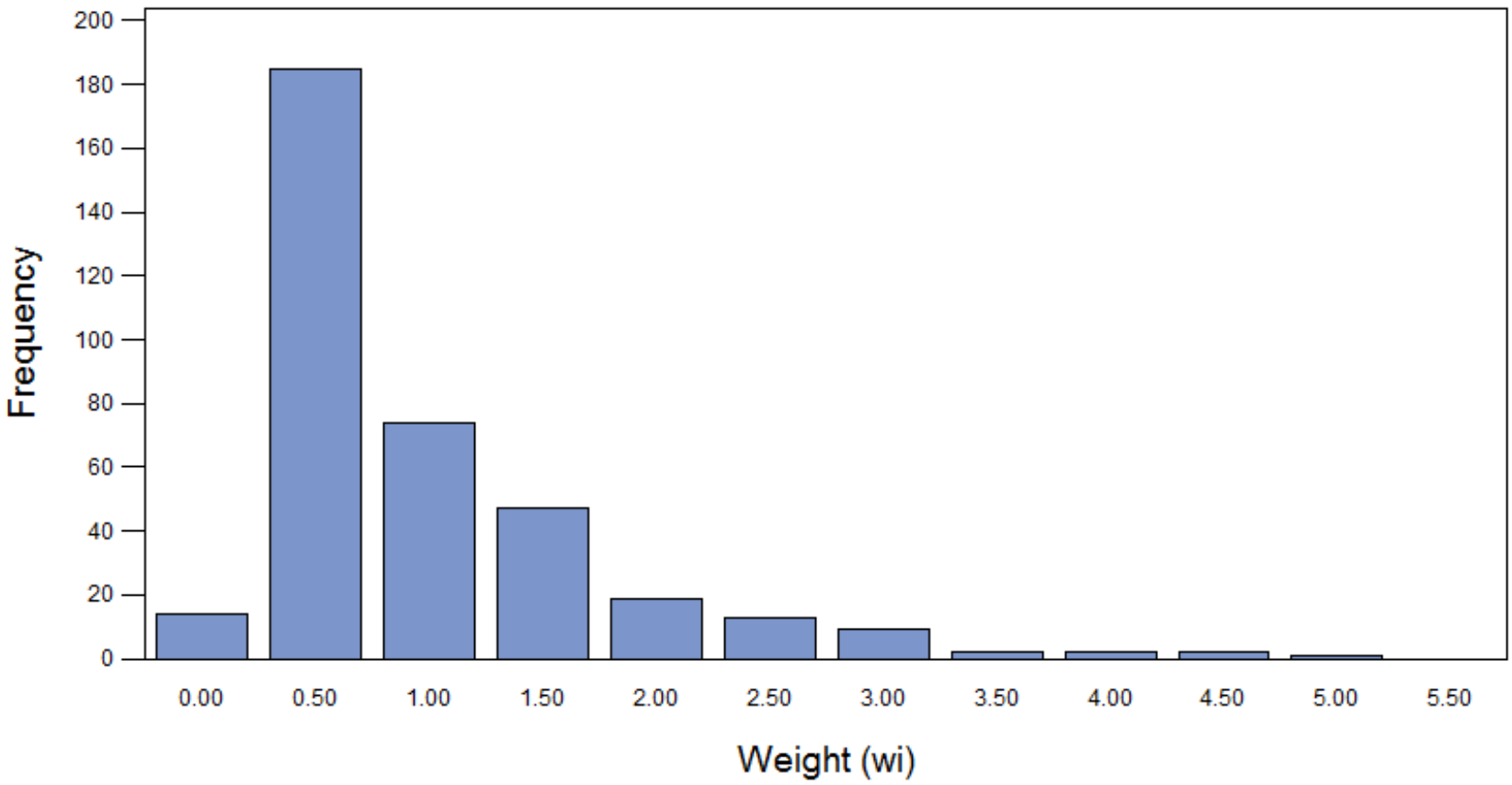
numerus
Discover Statistical Power

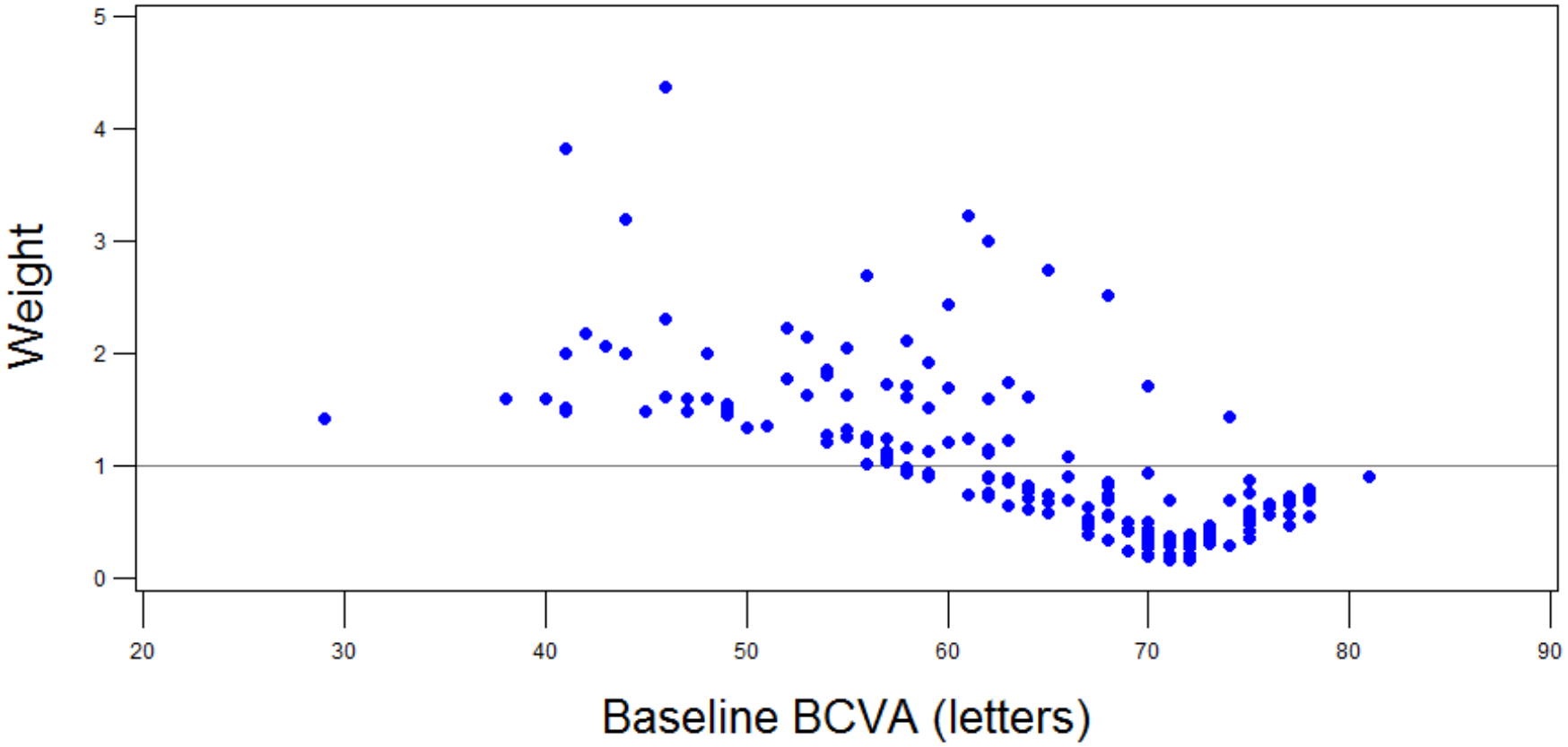
Source	Baseline BCVA (letters)		Baseline CRT (µm)	
	Laser ***Mean (SD)***	**Anti-VEGF** ***Mean (SD)***	**Laser** ***Mean (SD)***	**Anti-VEGF** ***Mean (SD)***
VIVID/ VISTA ***(target AGR)***	60.2 (10.8)	59.1 (11.0)	509 (153)	497 (151)
RESTORE/ RESPOND*	59.8 (9.8)	61.5 (9.5)	442 (133)	449 (126)
RESTORE/ RESPOND (weighted IPD)	*60.2 (10.8)*	*59.1 (11.0)*	*509 (153)*	*497 (151)*

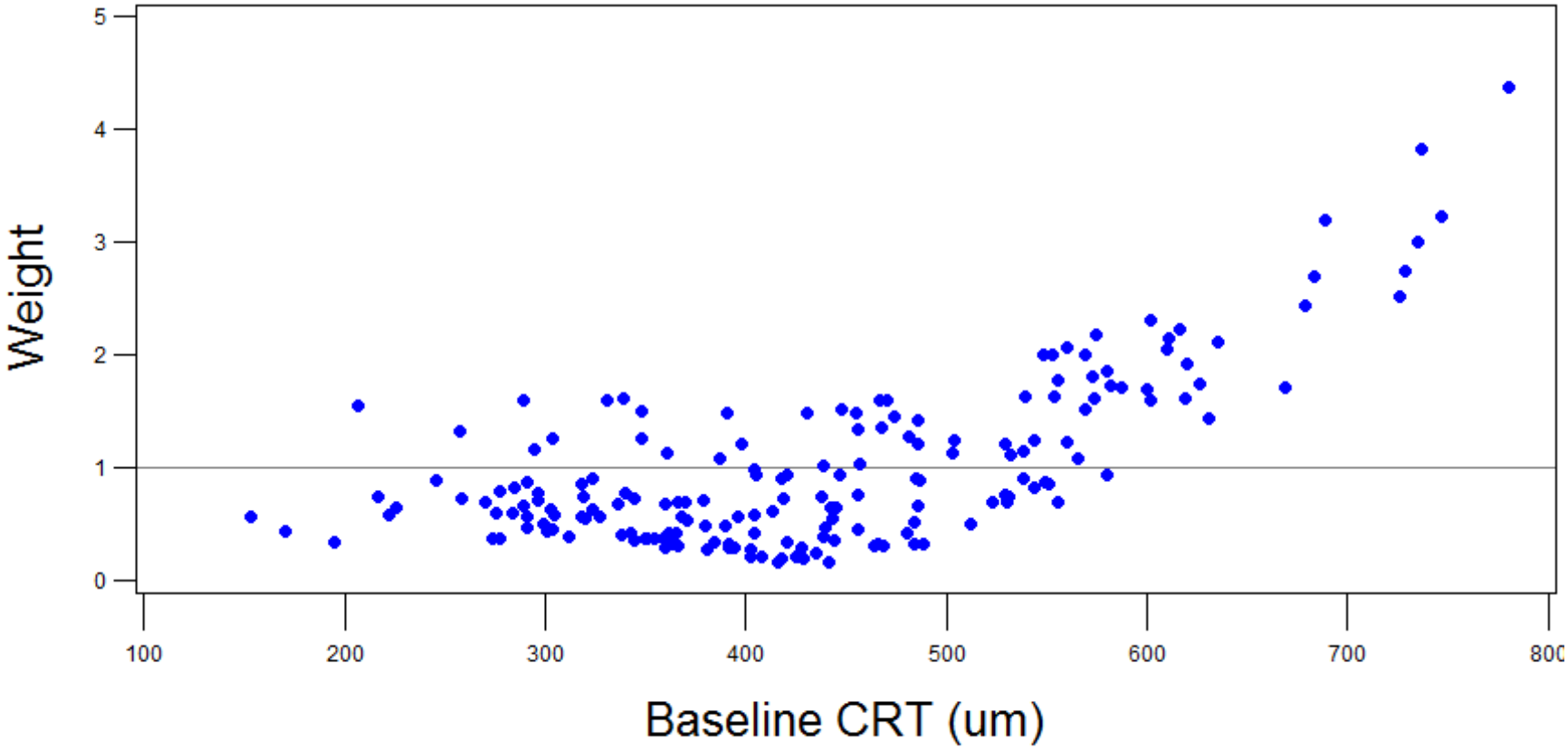
*VIVID/VISTA inclusion/exclusion criteria applied.

Weights

numerus
Discover Statistical Power




Weight
5
4
3
2
1
0
20
30
40
50
60
70
80
90
Baseline BCVA (letters)


Weight
5
4
3
2
1
0
100
200
300
400
500
600
700
800
Baseline CRT (um)

Effect on BCVA Change to M12

Source	Anti-VEGF vs. Laser Mean (95% CI)
VIVID/VISTA	10.0 (8.3, 11.7)
RESTORE/RESPOND (raw IPD)	7.6 (5.3, 9.8)
RESTORE/RESPOND (weighted IPD)	10.3 (7.9, 12.7)

 |

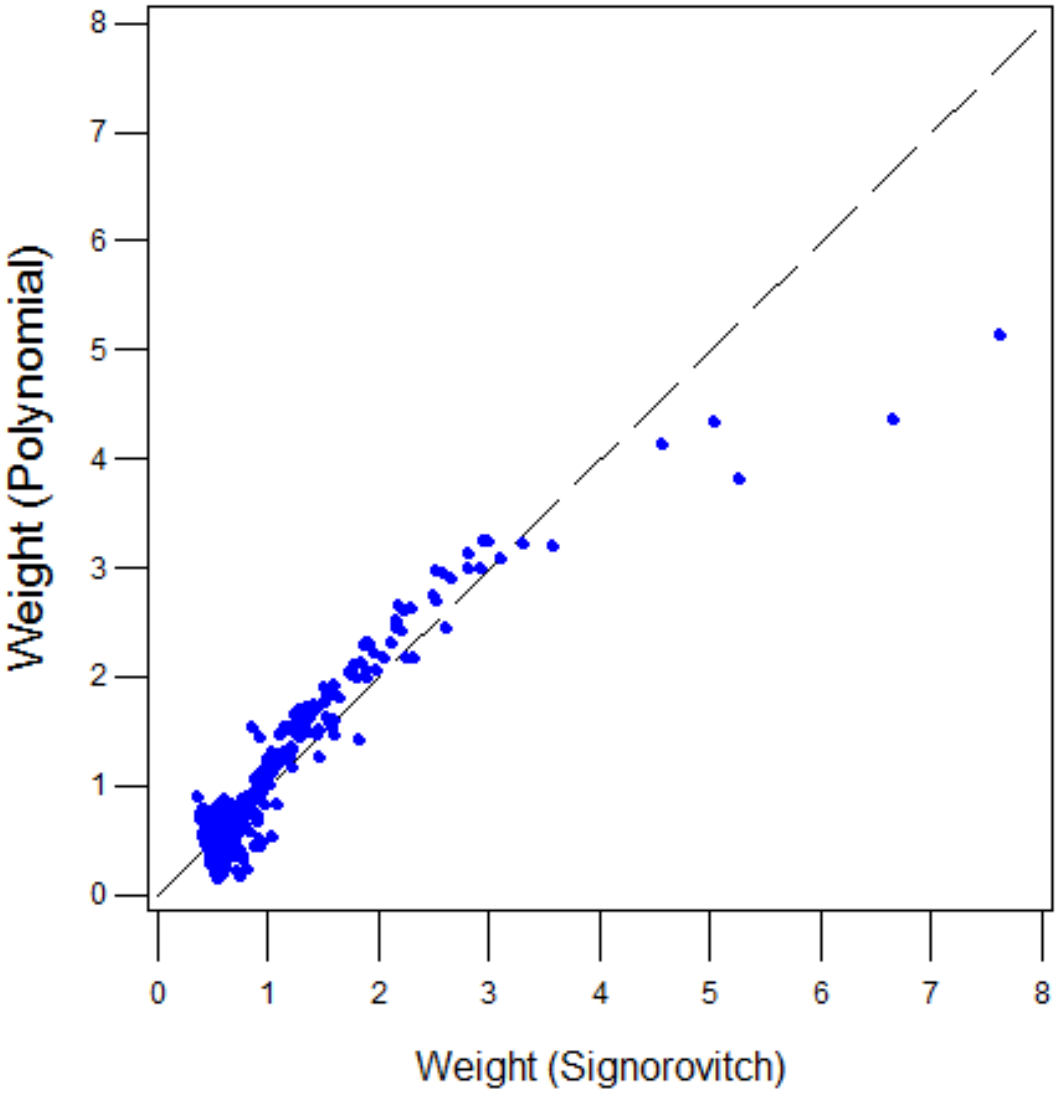
SAS Code (brute force approach)

numerus
Discover Statistical Power

```
* Make multiple copies of original dataset;
proc surveyselect data=d1 out=d2 method=srs sampsize = 368 reps=50000 seed=1;
run;

* Simulate beta weighting parameters for each baseline variable;
data d3;
 do Replicate=1 to 50000;
  a1=ranuni(0); b11=ranuni(0); b12=ranuni(0);b13=ranuni(0);b14=ranuni(0);*BCVA;
  a2=ranuni(0); b21=ranuni(0); b22=ranuni(0);b23=ranuni(0);b24=ranuni(0); *CRT;
 output;end;
run;

* Calculate weight;
data d4;
 merge d2 d3;
 by Replicate;
 t1=(BCVA-min1)/(max1-min1);
 t2=(CRT  -min2)/(max2-min2);
 w1=abs(b11*(t1-a1)+b12*(t1-a1)**2+b13*(t1-a1)**3+b14*(t1-a1)**4);
 w2=abs(b21*(t2-a2)+b22*(t2-a2)**2+b23*(t2-a2)**3+b24*(t2-a2)**4);
 wo=w1+w2;* Overall weight;
run;
```

```
* Calculate weighted means and SDs;
proc means data=d5 noprint;
 var BCVA CRT;
 weight wo
 by Replicate trt;
 output out=ss2 mean=m1 m2 std=s1 s2;
run;

* Compare vs. AGR: target means / SDs , loss function;
data ss3;
 set ss2;
 if trt='Ranibizumab' then do;mean1=59.1;std1=11.0;mean2=497;std2=151; end;
 if trt='Laser'       then do;mean1=60.2;std1=10.8;mean2=509;std2=153; end;
 dm1=(m1-mean1)**2;
 dm2=(m2-mean2)**2;
 sm1=(s1-std1)**2;
 sm2=(s2-std2)**2;
 r1=sqrt(dm1)/std1; r2=sqrt(sm1)/std1;
 r3=sqrt(dm2)/std2; r4=sqrt(sm2)/std2;
 loss=r1+r2+r3+r4; * Loss function;
run;
```

Polynomial weighting vs. Signorovitch



Effective Sample Size:

N= 368
Signorovitch = 221
Polynomial = 229

Polynomial Weighting: Weaknesses

numerus
Discover Statistical Power

- Cannot match against what is not reported
- Lack of distribution overlap between IPD and AGR
- Weights are estimates (bootstrap CIs?)
- How to choose maximum weight?
- How to maximise ESS?
- Best trade off between matching AGR & max ESS?

Matching vs. Effective Sample Size

numerus
Discover Statistical Power

IPD: 1, 2, 6, 7

AGR target: mean=4, SD=2.94

How to weight data, i.e. choose w_1, w_2, w_3, w_4?

Simple(!) solution: $w_1=w_2=w_3=w_4=1$ (ESS=4.0)

Suboptimal solution: $w_1=0.57$, $w_2=1.64$, $w_3=0.37$, $w_4=1.42$ (ESS=3.1)

Polynomial Weighting: Benefits

numerus
Discover Statistical Power

- ‘Convergence’ appears to be good

- Allows maximum weights to be specified
- Flexible (user-defined) loss function:
  - weight some baseline characteristics more heavily than others
  - weight some summary stats more others (e.g. mean > SD)
  - match against multiple summary stats (e.g. mean, median, SD, proportions, IQR, other percentiles)
  - could incorporate ESS maximisation

jonathan.alsop@numerus.com

numerus

Discover Statistical Power

www.numerus.com |

MAIC: Software

numerus
Discover Statistical Power

Signorovitch

- SAS: NLPNRA subroutine and GRD option in PROC IML
- R: Optim function

Malangone and Sherman

- SAS

Polynomial weighting

- SAS: Brute force, Proc IML?
- R : Optim function