Statistical Considerations in Multi-Regional Clinical Trials

Yoko Tanaka

Eli Lilly and Company

yokot@lilly.com

BASS XVII November 8th, 2010

Disclaimer

- The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and are not necessarily of the Eli Lilly and Company.

Outline

- MRCT - Challenges & Opportunities
- Regulatory Guidance – US perspectives, ICH E5, Japan, China, Taiwan, Korea, EU
- Overall Consistency vs. Region/Country-Specific Consistency
  - Q1: Overall Consistency cross Regions?
  - Q2: For Individual Region, Are The Results Consistent With Overall Conclusions / References ?
- Definition of Region
- Case Example
- Discussion
- Summary

MRCT - Challenges & Opportunities

- **FDA/PhRMA Workshop:**
  Challenges & Opportunities of Multiregional Clinical Trials, Bethesda, MD October 29-30, 2007
  - The workshop brought together representatives from industry, regulatory agencies and academia, who shared their experiences, perspectives and expertise

  Opportunities: Worldwide registration (Faster access to new medication throughout the world), Diverse patient population, Cost efficiency,

  Challenges: Different health authority requirements, Different intrinsic /extrinsic factors

Regional Differences ?

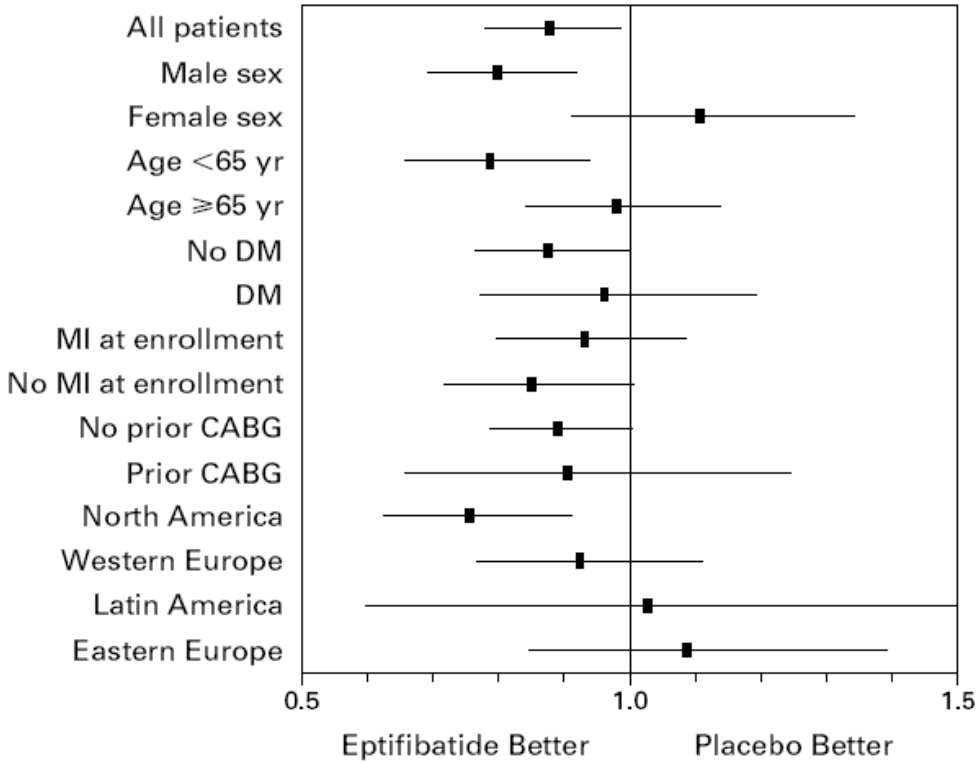
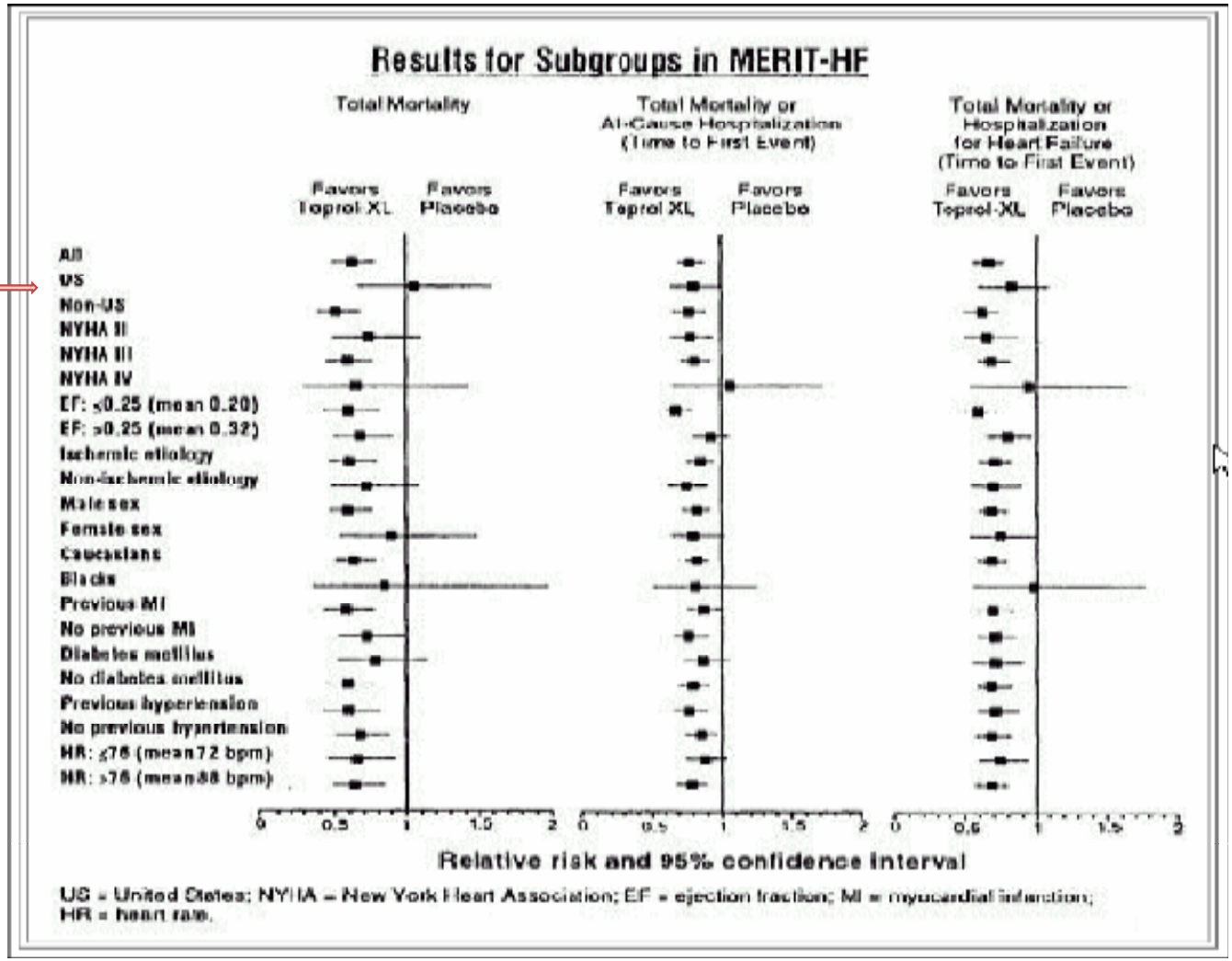


Figure 2. Odds Ratios for Death or Nonfatal Myocardial Infarction in Selected Subgroups of Patients.
The horizontal lines indicate 95 percent confidence intervals. DM denotes diabetes mellitus, MI myocardial infarction, and CABG coronary-artery bypass graft surgery.

THE PURSUIT trial. Inhibition of platelet glycoprotein IIb/IIIa with eptifibatide in patients with acute coronary syndromes. N Engl J Med 1998;339:436-43.

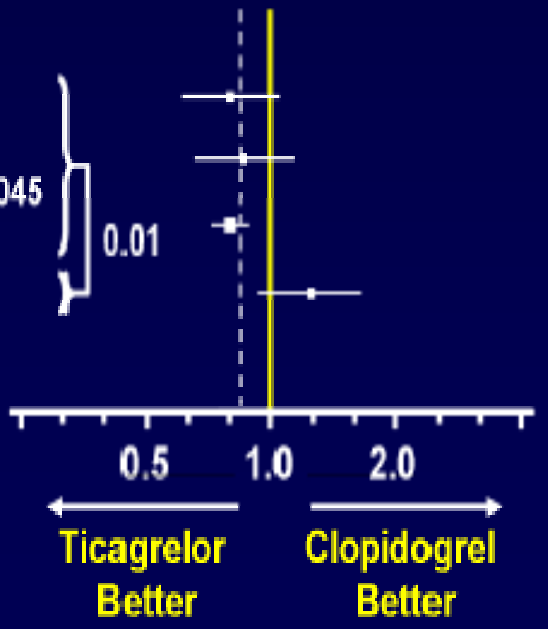
Regional Differences ?



From US Label

Regional Differences ?

Characteristic	Total Patients	# of Events Tic	# of Events Clop	HR (95% CI)	Interaction p-value
Geographic Region					
Asia / Australia	1714	95	116	0.80 (0.61, 1.04)	
Cent / Sth America	1237	91	104	0.86 (0.65, 1.13)	
Euro / Md E / Afr	13859	576	712	0.80 (0.72, 0.90)	0.045
North America	1814	102	82	1.25 (0.93, 1.67)	0.01



Source: Sponsor presentation at CV and Renal Drugs Ad Comm Meeting July, 2010 CC-30

Lesson

To date, no or very little consideration is given for region assessment and consistency assessment in the design stage.

We need to better plan for heterogeneity of regional differences.

Regulatory Guidance ?

- US Perspectives
- ICH E5 (1998) – Ethnic Factors in the Acceptability of Foreign Clinical Data
- Japan - Basic principles on Global Clinical Trials (2007)
- China – SFDA Guidance
- Taiwan – Double Twelve Announcement - Bridging Guidance (2004)
- Korea – KFDA, Guidelines for an Evaluation of Bridging Data (2008)
- EU - Reflection Paper on The Extrapolation of Results From Clinical Studies Conducted Outside Europe to The EU Population (2009)

From US FDA's perspectives

- Recent years, multi-national trials include Latin American, Eastern European and Asian sites, often with U.S. and W. Europe representing a minority of patients.

- We accept this, but we are looking closely at regional differences and have a degree of nervousness, stimulated by a number of troubling examples

Bob Temple, PhRMA MRCT Workshop, 2007

From US FDA's perspectives

- For every multi-regional study, create a common template for planning for homogeneity/heterogeneity of regional differences and exploring sample sizing according to assumptions of dropouts, follow-up, compliance, event ascertainment by investigator, degree of internal consistency
- Improve the statistical analysis plan to specifically address strategies and interpretation of heterogeneity, power, internal consistency of by region results
- Plan and size the trial for expected heterogeneity - must decide what type and how much (metrics of consistency - qualitative vs. quantitative)

Bob O'Neil, 2009 FDA/Industry Statistics Workshop

ICH E5

Guidance for Industry

E5 – Ethnic Factors in the Acceptability of Foreign Clinical Data

---

Questions and Answers

ICH: International Conference on Harmonisation

ICH E5 – Intrinsic/Extrinsic Factors

APPENDIX A

Classification of intrinsic and extrinsic ethnic factors

INTRINSIC		EXTRINSIC
Genetic	Physiological and pathological conditions	Environmental
Gender	Age (children-elderly)	Climate
		Sunlight
Height		Pollution
Bodyweight		
	Liver	**Culture**
	Kidney	Socioeconomic factors
	Cardiovascular functions	Educational status
ADME		Language
Receptor sensitivity		
Race		Medical practice
		Disease definition/Diagnostic
Genetic polymorphism of the drug metabolism		Therapeutic approach
		Drug compliance
		Smoking
		Alcohol
		Food habits
Genetic diseases	Diseases	Stress
		Regulatory practice/GCP
		Methodology/Endpoints

Consistency from ICH E5 Q&A #11

- A multi-regional trial for the purpose of bridging could be conducted in the context of a global development program designed for near simultaneous world-wide registration. The objectives of such a study would be: **1) to show that the drug is effective in the region and 2) to compare the results of the study between the regions with the intent of establishing that the drug is not sensitive to ethnic factors.** The primary endpoint(s) of the study should be defined and acceptable to the individual regions and data on all primary endpoints should be collected in all regions under a common protocol. In instances where the primary endpoints to be used by the regions are different, data for comparison purposes on all primary endpoints should be collected in all regions.

Japan guidance - Basic principles on Global Clinical Trials (2007)

- Q1. Basic requirements ?
- Q2. Timing ?
- Q3. A necessity of Phase I or PK information in Japanese ?
- Q4. A necessity of a dose-finding study in Japanese ?
- Q5. Basic points in designing
- **Q6. Sample size**
- Q7. Overseas Endpoint
- Q8. Appropriateness to conduct a identical but smaller trial to the global trial in Japan separately
- Q9. Requirement for a control group ?
- Q10. How to define concomitant medications or therapies
- Q11. Desirable disease area
- Q12. A flow-chart for conducting a global trial

http://pmda.go.jp/oprations/notice/2007/file/0928010-e.pdf

Basic Principles on Global Clinical Trials (1997)

- Question 6:
  - "To conduct an appropriate global clinical trial, how should the sample size and the proportion of Japanese subjects be determined?"
- Answer 6:
  - "… a global trial should be designed so that consistent results can be obtained between results from the entire population and the Japanese population."
  - Two examples for sample size estimation based on point estimates (Method 1 and Method 2)

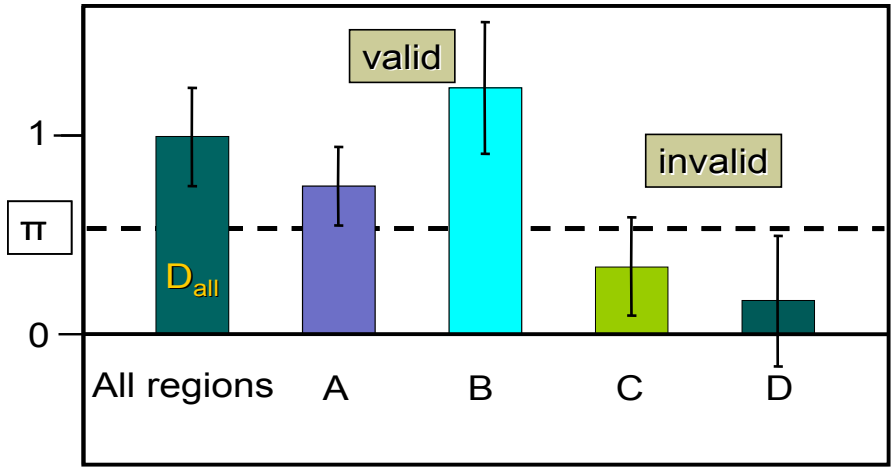
Japanese Guidance - Method 1

Estimated treatment effect D_{all} and D_J for all patients and Japanese cohort, respectively. Results are considered consistent if { $D_J / D_{all} > \pi$}.

Determine the number of Japanese subjects so that

- $D_J / D_{all} > \pi$ will occur with a probability of 80 % or higher.
- A π should be set as an appropriate value, and 0.5 or more is generally recommended. When this method is used, the following relationships will be observed: An attempt to minimize the Japanese sample size will increase the total sample size, and an attempt to minimize the total sample size will increase the Japanese sample size.

Subject numbers from Japan should be enough to assure "$D_J / D_{all} > \pi=0.5$"

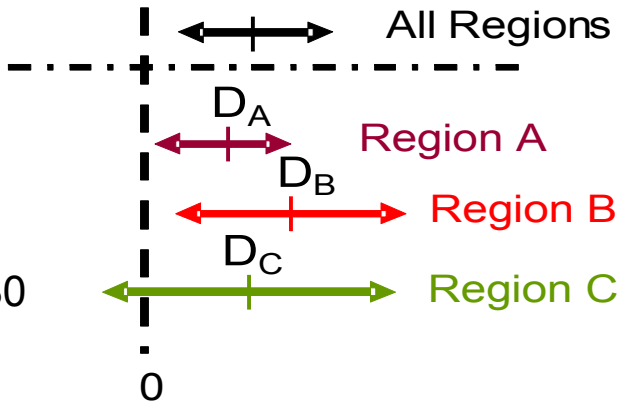


Japanese Guidance - Method 2

Positive trend in all regions. Estimated treatment effect D_{all} for all patients and D_k , k=1, …, *k* for region *k*, respectively. Results are considered consistent if { $D_k > 0$, k=1, …, *k* }.

- Number of subjects is determined so that each of the D_k will exceed 0 with a probability of 80 % or higher.
- The probability tends to increase if equal number of subjects is enrolled from each region. This method allows considering the Japanese sample size without changing the total sample size. However, it should be noted in this method that sufficient interregional comparison may not be possible when Japanese component ratio is small and number of subjects are few.

Each of the D_A, D_B, and D_C will exceed 0 with a probability of 80 % or higher



$D_A>0$ & $D_B>0$ & $D_C>0$

China Guidance to Meet Local Regulatory Requirements

- Global IMCT: IMCT conducted for the purpose of first approval in US/EU.
- Regional IMCT: IMCT conducted with at least other 2 countries, mostly in Asia Pacific region, after the initiation of global IMCT but before the CPP is available in US/EU.
- Local trial: Stand-alone China study after CPP is available in US/EU.

IMCT: International Multi-center Clinical trial
CPP: Certificate of Pharmaceutical Product, e.g, marketing license in US/EU

Statistical Significance of China data for approval?

- Global IMCT: No.

  100 pts per arm required

  SFDA will look at trial level statistical significance and the trend of China data.

- Regional IMCT: recommended

  In most cases, 100 pts per arm + statistical significance of China data recommended.

  but flexibility allowed if unreasonably large sample size required, say more than 600 pts in China population.

- Local trial: Yes.

  100 pts per arm + statistical significance of China data required

Additional documents needed for China submission

- China specific CSR (clinical study report): required. This means the analysis on the whole population need to be repeated in Chinese patients.
- China specific SAR (statistical analysis report): required.
- Asian CSR (or ethnic sensitivity report): optional.
- Clinical datasets (Chinese patients only): required
- All documents need to be translated into Chinese: required.

Bridging Guidance Documents for Taiwan, Korea, EU

- Taiwan :
  - A Guidance of Bridging Study “Ethnic Factors in the Acceptability of Foreign Clinical Data” - May, 2002.
- Korea:
  - Guidelines for an evaluation of bridging data (2008) :
    - Guidelines for the evaluation of foreign clinical data and bridging data and for the decision of bridging study
- EU:
  - Reflection Paper on The Extrapolation of Results From Clinical Studies Conducted Outside Europe to The EU-Population (2009)

Multi-Regional Clinical Trials

CONSISTENCY ASSESSMENT AND SAMPLE SIZE CONSIDERATION

Deeper Look on Consistency Criteria / Assessment

- Quan et al. Assessment of consistency of treatment effects in multiregional clinical trials. *Drug Information Journal 2010; 44:617–632.*

- Chen et al. Assessing consistent treatment effect in a multi-regional clinical trial: a systematic review. *Pharmaceutical Statistics 2010;9:242–253.*

These papers are by PhRMA MRCT K ey Issue Team – consistency assessment workstream

Overall Consistency vs. Region/Country-Specific Consistency

Q1: Overall consistency across **regions**?

- To support the robustness of the conclusions
- All regions are considered equal
- Should be pre-specified in the Protocol

Q2: For individual **region**, are the results consistent with overall conclusions / references ?

- Frequently requested by local regulatory agency to support registration; sometimes at the country level
- Cautionary about subgroup analysis (e.g., multiplicity)
- Sometime not possible to pre-specify in the Protocol

How to Define Region

- Region does not necessarily have to be geographical or political.
- Different factors can be considered depending on therapeutic area / disease state.
- **"Region" should be pre-defined and how it will be analyzed should be pre-specified in the planning stage (stratification, consistency method should be integral in the design).**

2010 DIA MRCT Workshop – Session 4B, Definition of Region

Definition of Region – Factors & Rationale to Consider

Factors	Rationale
Race and ethnicity	Surrogates for genomic issues and therefore a supposedly homogeneous, w.r.t. drug effects, group
Medical practice	Encompasses practices of treating a patient including local medicines, hospital treatment
Human Development Index	Surrogate for ability to provide and have access to the "latest" developments in health care (Adult literacy, GDP, Education, life expectancy)
Disease Epidemiology	Goes to the differing characteristics of the disease (including genomics/biomarkers) which are reflected by many of the issues on this list. Provides the background information that can indicate where disparate characteristics occur that will affect the planning, analysis and execution of the clinical trial.
Geographic proximity	The traditional idea of a region, yet still very fluid
Geopolitical / Institutional	Health Authority driven
Culture	Broad term to encompass common health practices, ethics, and behaviors that impact on a clinical trial that arise from within a common culture.

Overall Consistency Across Regions

- Consistency Definition:
  Absence of **treatment-by-region interaction**

$$H_0^4 : \delta_1 = \delta_2 = ... = \delta_s = \delta$$

- Parametric approaches
  - Testing treatment-by-region in the regression models
  - Can also adjust for baseline prognostic factors
  - Fixed vs. random effects model
- Non-parametric homogeneity tests
  - Cochran's Q, Breslow-Day for odds ratios, etc

- Considerations:
  - At significance level 0.10: inflated false positive rate?
  - Focus on regional deviations from the overall treatment effect, rather than the actual effect in individual regions
  - Quantitative vs. qualitative interactions (e.g., Gail-Simon)

Overall Consistency Across Regions
Quantifying Heterogeneity: I^2

- Testing hypothesis vs. quantifying heterogeneity
  - Statistical vs. clinical significance: the p value does not reasonably describe the extent of heterogeneity, which is more relevant clinically

- Higgins et al defines, $I^2 = 100\% \times (Q - \text{df})/Q$
  - Cochran's Q statistic ~ Chi-Square with degrees of freedom (df) = s-1, s = #regions
  - I^2 between 0% -100% (assigned 0 if negative)
  - % total variation across studies due to heterogeneity
    - $I^2 = 0\%$: no observed heterogeneity
    - Larger values show increasing heterogeneity
  - I^2 value of 25%, 50%, 75% considered low, moderate, high heterogeneity, respectively

Overall Consistency Across Regions

- Consistency Definition:
  Absence of **significant difference for any regions from overall**

- One-sided version of interaction definition:

$$H_0^5 : \delta_i \geq \delta$$

- Accept all $H_{0i}'s$ if all

$$Z_i^* = \hat{\delta}_i - \hat{\delta} + z_{\alpha'}\sigma\sqrt{\frac{2}{N}(\frac{1}{f_i} - 1)} > 0$$

Individual Region Consistency

- **Method 1**. Estimated treatment effect $\hat{\delta}$ and $\hat{\delta}_{\mathrm{J}}$ for all patients and Japanese cohort, respectively. Results are considered consistent if { $\hat{\delta}_{\mathrm{J}} / \hat{\delta} > \pi$ }, where π =0.5 or higher

- **Method 2** (Positive trend across all regions) Estimated treatment effect $\hat{\delta}_i$ for Region *i* (*i*=1, …, *s* regions). Results are considered consistent if {$\hat{\delta}_i$ > 0, *i*=1, …, *s*}

- PhRMA work on variations of Method 1 and Method 2.

Individual Region Consistency

Definition 1 - Exceeding a proportion of the observed overall effect

- Combine Method 1 and Method 2 of the Japanese Guidance

$$\hat{\delta}_1 > \pi\hat{\delta}, \hat{\delta}_2 > \pi\hat{\delta}, ..., \hat{\delta}_s > \pi\hat{\delta} \quad (*)$$

where π could depend on *s*. A special case of π=0 is Method 2 in the guidance.

- When $N_1 = N_2 = N/2$ and π=0.5, (*) implies two-sided definition

$$1.5\hat{\delta} > \hat{\delta}_1 > 0.5\hat{\delta}, 1.5\hat{\delta} > \hat{\delta}_2 > 0.5\hat{\delta}$$

32

Individual Region Consistency

Definition 2 - Exceeding a constant effect size

- For a constant *b,* consistency is defined as*:*

$$\hat{\delta}_1 > b, \hat{\delta}_2 > b, \ldots, \hat{\delta}_s > b$$

Method 2 of Japanese guidance is a special case (*b*=0)

- Case (1): $\hat{\delta}_1 = 15, \hat{\delta}_2 = 5, \hat{\delta} = 20/2 = 10$

- Case (2): $\hat{\delta}_1 = 7.5, \hat{\delta}_2 = 3.5, \hat{\delta} = 11/2 = 5.5$

- Effects for all regions are better for (1) than (2)
- Consistency for (2) but not for (1) based on Def. 1 (π=0.5)
- Consistency for (1) but not for (2) based on Def 2. if *b*=4.5

Definition 3 - Exceeding a prop of the observed overall effect using hypothesis testing

- To reject

$$H_0^3 : \delta_1 \leq \pi\delta \text{ or, ... } \delta_s \leq \pi\delta$$

and accept

$$H_a^3 : \delta_1 > \pi\delta \text{ and, ... } \delta_s > \pi\delta$$

- Even if π =0, there is no power unless the significance level α' is large. When α'=0.5, Def 3 becomes Def 1.

Other Possible Methods

- Two-sample t-test H_0: $\mu_1 = \mu_2$
- Bioequivalence $[-\theta_A , \theta_A]$
- Non-inferiority H_0: $\delta_i < \pi\delta$
- Overlapping Coefficient / Proportion of Similar Response (PSR)
- Shih's Consistency Criteria for bridging study

Unconditional & Conditional Power

- Unconditional power for showing consistency:

$$\Pr(\hat{\delta}_i > \pi\hat{\delta}, i = 1,\ldots,s \mid \delta_i, i = 1,\ldots,s)$$

- Consistency assessment is meaningful only if the overall treatment effect is significant.
- Conditional power:

$$\Pr(\hat{\delta}_i > \pi\hat{\delta}, i = 1,\ldots,s \mid \hat{\delta} - z_\alpha \sigma\sqrt{2/N} > 0)$$

Let f_i denote the proportion of study patients within region i, and u_i the ratio of the treatment effect in region i to the overall effect.

Simulation Results - Unconditional and conditional power for claiming consistency (defs 1-3)

Definitions 1-3: s=3 ($\alpha = 0.025$, $\delta = 0.25$, $\sigma = 1$)

(f_1, f_2, f_3)	(u_1, u_2, u_3)	Uncond.	Cond.	Uncond.	Cond.	Uncond.	Cond.
$1-\beta$ =0.8		Definition 1 π =1/3		Definition 2 b=0.083		Definition3 π =0, α' =0.3	
(1/3,1/3,1/3)	(1,1,1)	67	76	64	76	64	77
(0.2,0.2,0.6)	(1,1,1)	62	69	59	70	56	66
(1/3,1/3,1/3)	(0.9,1,1.1)	66	75	63	75	64	76
(1/3,1/3,1/3)	(0.6,1.2,1.2)	59	67	57	68	57	69
(0.2,0.2,0.6)	(0.7, 0.7, 1.2)	45	51	45	53	40	48
(0.2,0.2,0.6)	(1.2, 1.1, 0.9)	68	76	64	76	62	73
(0.2,0.4,0.4)	(0.8,1.1,1)	61	69	58	69	57	67
(0.1,0.45,0.45)	(1.9,0.9,0.9)	72	81	67	80	68	80

Definitions 1-3: s=4 ($\alpha = 0.025$, $\delta = 0.25$, $\sigma = 1$)

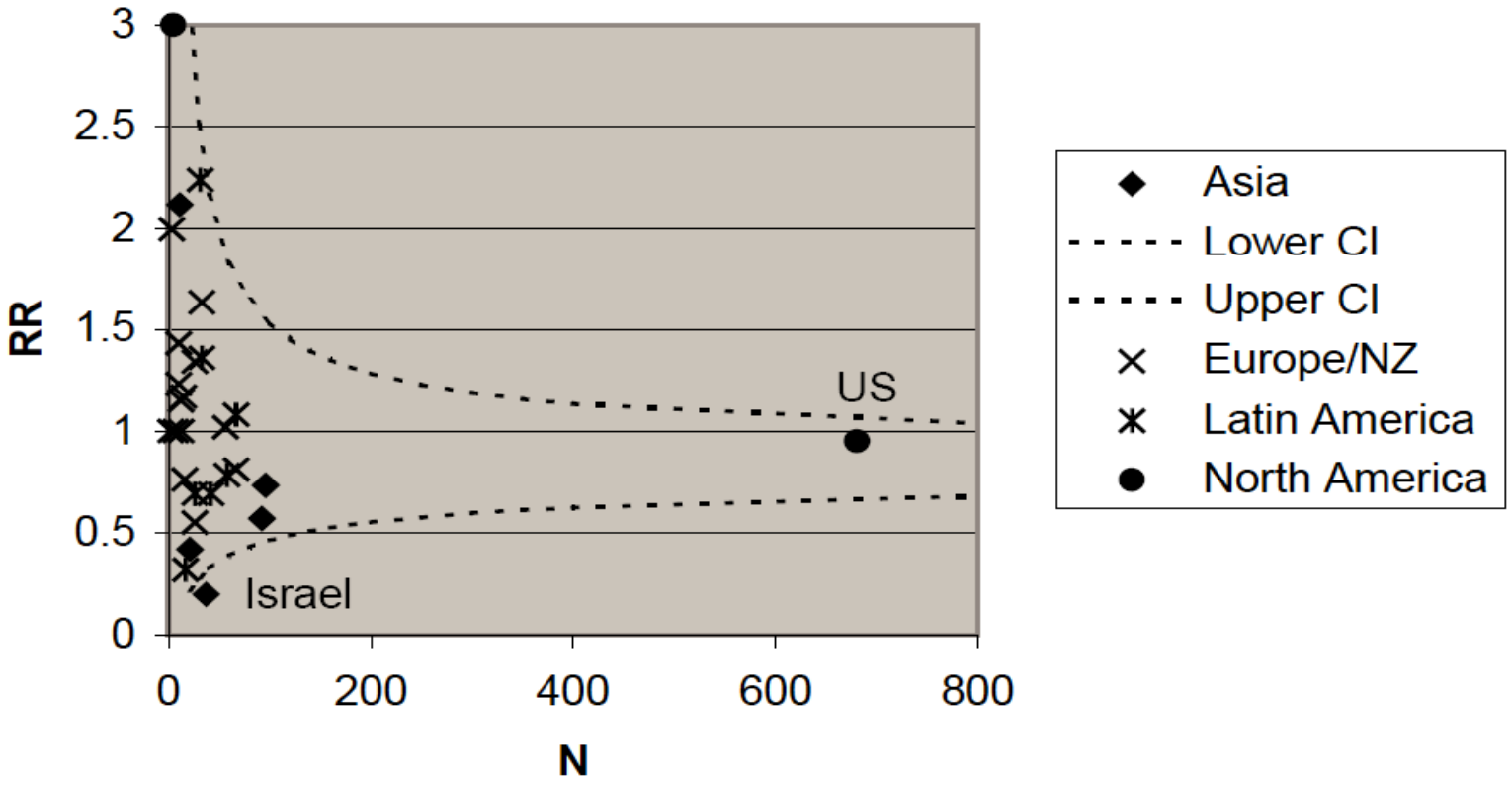
(f_1, f_2, f_3, f_4)	(u_1, u_2, u_3, u_4)	Uncond.	Cond.	Uncond.	Cond.	Uncond.	Cond.
$1-\beta = 0.8$		Definition 1 $\pi = 1/4$		Definition 2 $b = 0.063$		Definition 3 $\pi = 0$, $\alpha' = 0.4$	
(1/4,1/4,1/4,1/4)	(1,1,1,1)	54	64	53	64	59	70
(0.1,0.3,0.3, 0.3)	(1,1,1,1)	51	60	50	60	54	64
(1/4,1/4,1/4,1/4)	(0.7,0.8,0.8,1.7)	44	52	44	53	49	59
(1/4,1/4,1/4,1/4)	(0.7,1.1,1.1,1.1)	52	61	51	61	56	67
(0.1,0.3,0.3,0.3)	(1.3,1,0.9,1)	54	63	53	64	58	68
(0.1,0.3,0.3,0.3)	(1.9,0.9,0.9,0.9)	57	66	55	67	61	72
(0.15,0.15,0.35,0.35)	(0.65,0.65,1.1,1.2)	39	45	39	47	41	49
(0.15,0.15,0.15,0.55)	(1.3,1.4,1.4,0.7)	60	69	58	69	63	74

Simulation Results – Main Observations

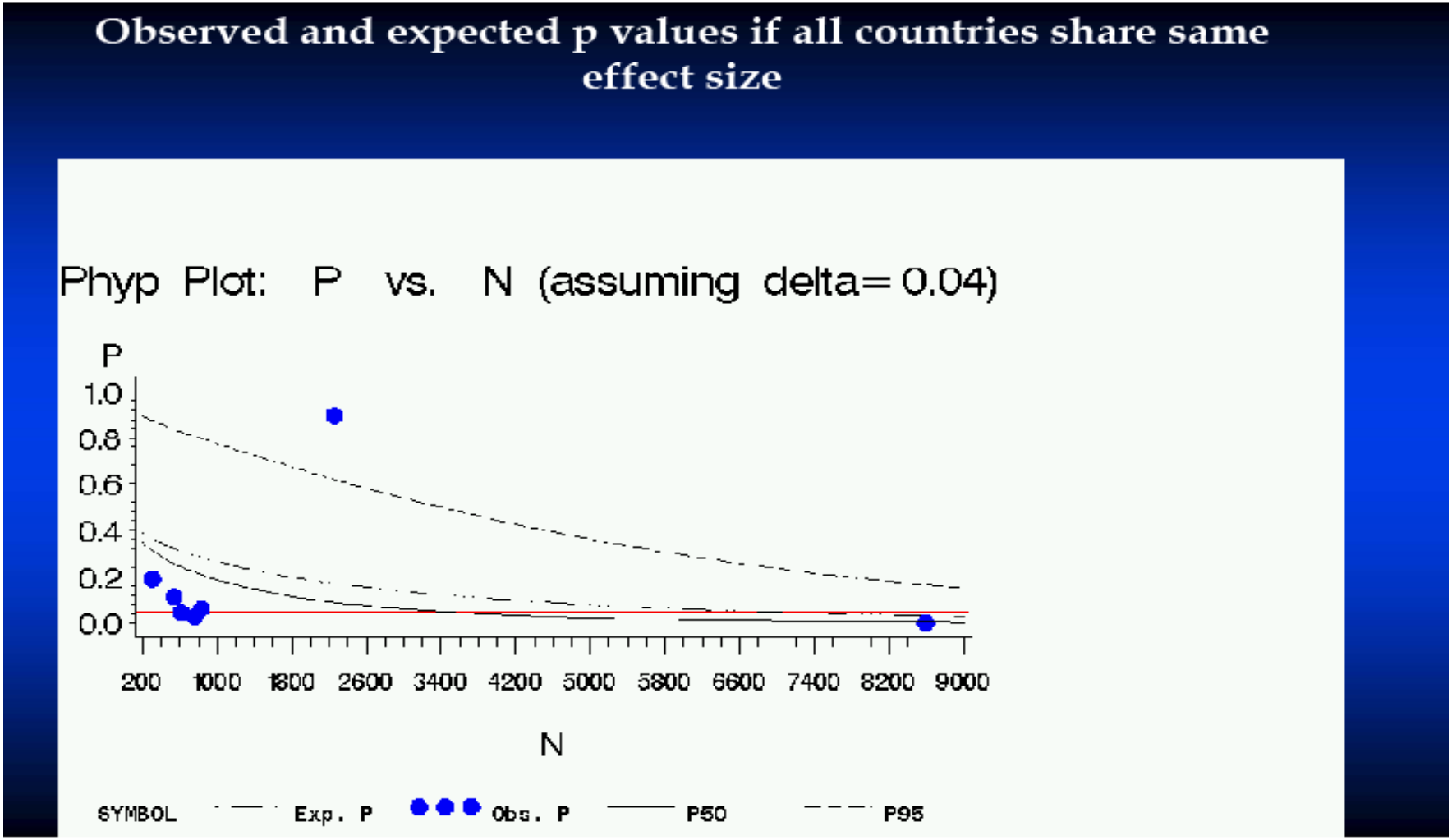
- Power higher with 3 regions compared 4.
- Similar power among definitions 1-3.
- Sample size implication

Programs in R are available for sample size and power explorations.

Analysis Presentation – Funnel Plot

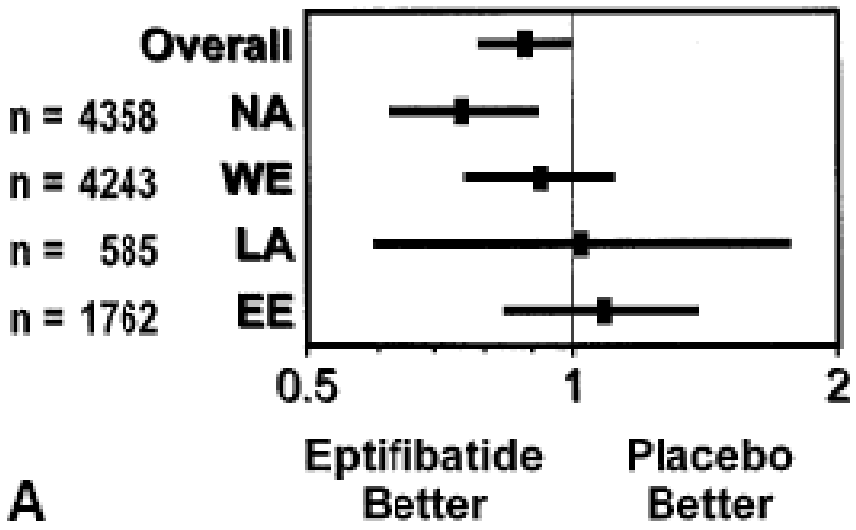


Analysis Presentation – Phyp Plot



PURSUIT – Case Example

- Eptifibatide vs Placebo
- Primary endpoint: death or MI within 30 days
- Overall event rates 14.2% (eptifibatide) vs. 15.7% (placebo); p=0.042
- No apparent treatment effect for Latin America and Eastern Europe



0.89 (0.79, 0.99)
0.75 (0.63, 0.91)
0.92 (0.77, 1.11)
1.03 (0.60, 1.76)
1.09 (0.85, 1.39)

Odds Ratios and 95% CIs
NA: North America
WE: Western Europe
LA: Latin America
EE: Eastern Europe

Overall Consistency Across Regions

- May consider a logistic regression model for interaction test if individual patient data available
- Alternatively, may consider nonparametric homogeneity tests

Homogeneity Test	Chi-Square Statistic	DF	P-value
Cochran's Q	6.64	3	0.08
Breslow-Day	6.72	3	0.08

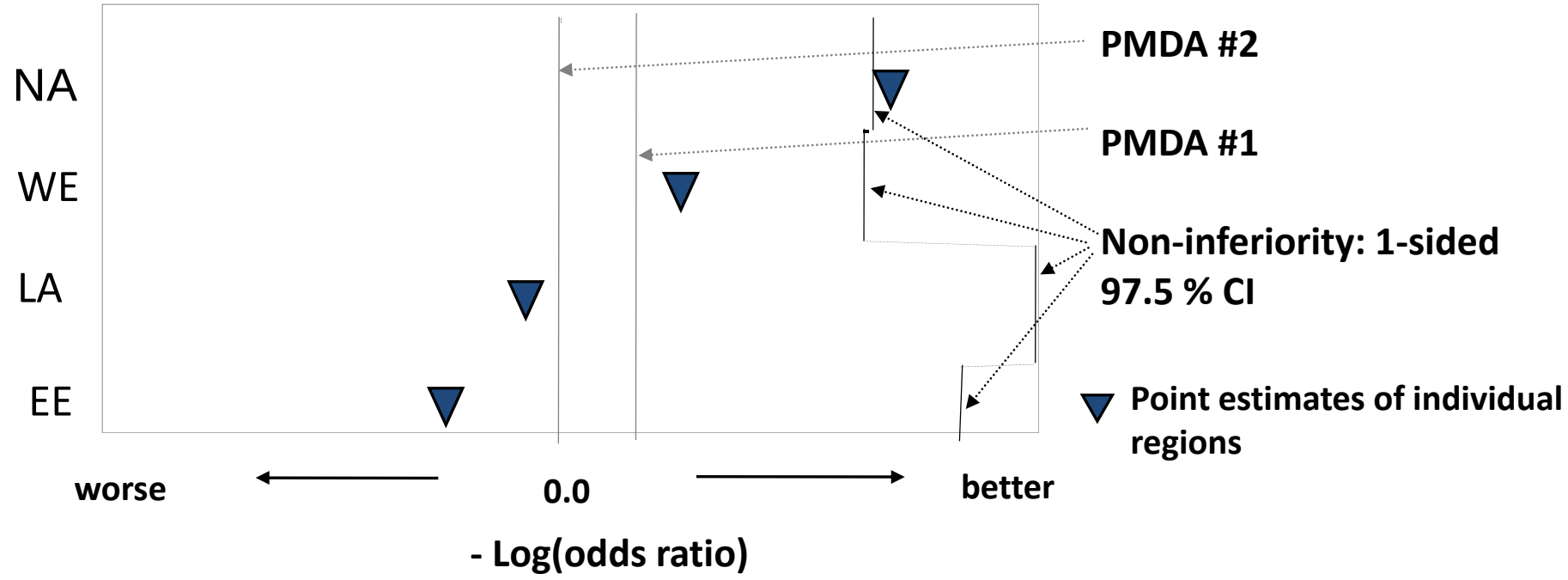
- Quantify the heterogeneity using Higgins I^2

  100% x (6.64-3)/6.64 = 55%, with 95% CI (0%, 85%)

I^2 value of 25%, 50%, 75% considered low, moderate, high heterogeneity

Regional Estimates

- PMDA #2: π =0
- PMDA #1 : π =0.5
- Non-inferiority : π =0.5



Summary – Case Study PURSUIT

- Eptifibatide was shown to be superior to placebo based upon overall data (OR=0.89 with 95% CI 0.79-0.99)
- A moderate heterogeneity across regions was observed
  - Higgins I^2 = 55%
- Interaction tests showed a borderline statistical significance
  - P=0.08 based upon Cochran's Q / Breslow-Day tests
  - Quantitative (vs. qualitative)
  - Not clear due to which region(s)
- Approach to assess consistency across regions based upon individual effect size indicated less effectiveness in LA and EE
  - Primarily because point estimates < 0

Discussion

- The two analyses try to address two different questions
  - ✓ Q1: Overall consistency across regions?
  - ✓ Q2: For individual region, are the results consistent with overall conclusions / references?

- Should we prospectively plan to answer Q2 ? OR should this be post-hoc ?
  - o Multiplicity issue due to subgroup analysis
  - o Interpretation of regional findings: exploring intrinsic/extrinsic factors upon unexpected findings
  - o Much smaller sample size in individual countries
  - o Possible signal detection instead of confirmatory

➢ **Consistency assessment warrants to evaluate key factors / trial conduct at the design stage (beyond statistical consideration).**

Discussion - Factors to Consider in Design

- Diagnosis of Target Disease/Condition – Definition of disease
- Medical Practice – Treatment guideline, Concomitant medications
- Patient Selection – May differ due to the conduct of the study
- Primary Endpoint Selection – Culturally sensitive ?
- Dosage of Active Treatment
- Global Standards
- Standard Education / Training
- Common Interpretation of protocol Instructions / definitions / criteria
- Quality measure in each region

Summary

- In a MRCT, regions to include and definition of regions should be evaluated from intrinsic / extrinsic factors.
- Planning of robustness and consistency of treatment effects among regions should be integral in the design.
- Some definitions are based on purely the observed effects and others based on testing of parameters.
- Pre-work on sample size and power comparison will help us to select one consistency definition for the protocol.
- Overall and regional sample sizes play key role in power.
- It is desirable to limit the number of regions and define regions in a sensible way for characterizing region treatment effects.

Original PhRMA MRCT WG (Oct 2008)
Consistency Assessment Workstream

Hui Quan (Sanofi-Aventis)	Josh Chen (Merck)
Mingyu Li (Celgene)	Paul Gallo (Novartis)
Bruce Binkowitz (Merck)	Ekopimo Ibia (Merck)
Yoko Tanaka (Eli Lilly)	Peter Ouyang (Celgene)
Xiaolong Luo (Celgene)	Gang Li (Johnson & Johnson)
Shailendra Menjoge (Boehringer Ingelheim)	Steven Talerico (Schering-Plough)
Kimitoshi Ikeda (Novartis)	

PhRMA MRCT KIT
Region Definition Workstream

Yoko Tanaka (Eli Lilly)	Carmen Mak (Merck)
Brent Burger (Cytel)	Josh Chen (Merck)
Bruce Binkowitz (Merck)	Ekopimo Ibia (Merck)
Mike Rabbia (Roche)	

49

References

ICH International Conference on Harmonization Tripartite Guidance E5. Ethnic Factor in the Acceptability of Foreign Data, The US Federal Register, 1998; Vol. 83, 31790-31796.

Ministry of Health, Labour and Welfare of Japan. Basic Concepts for Joint International Clinical Trials. September 28, 2007.

European Medicines Agency. Reflection paper on the extrapolation of results from clinical studies conducted outside Europe to the EU-population (draft). EMEA Doc. Ref. CHMP/EWP/692702/2008. February 2009. Available at http://www.emea.europa.eu/pdfs/human/ewp/69270208en.pdf

Quan H, Zhao PL, Zhang J, Roessner M, Aizawa K. Sample size considerations for Japanese patients in a multi-regional trial based on MHLW Guidance. *Pharmaceutical Statistics* 2009.

Kawai N, Chuang-Stein C, Komiyama O, Li Y. An approach to rationalize partitioning sample size into individual regions in a multiregional trial. *Drug Information Journal* 2007; 42:139-147.

Quan et al. Assessment of consistency of treatment effects in multiregional clinical trials. *Drug Information Journal 2010; 44:617–632.*

Chen et al. Assessing consistent treatment effect in a multi-regional clinical trial: a systematic review. *Pharmaceutical Statistics 2010;9:242–253.*

Hung et al. Consideration of regional difference in design and analysis of multi-regional trials. *Pharmaceutical Statistics 2010;9:173–178.*

Yang N, Li H, Criswell LA,Gregersen PK, Alarcoive diallelic Dn-Riquelme ME, Kittles R, Shigeta R, Silva G, Patel PI, Belmont JW, Seldin MF (2005) Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Human Genetics* 118, 382-392.

Halder I, Shriver M, Thomas, M Fernandez JR, and Frudakis T (2008) A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. *Human Mutation* 29, 6480658

References

Akkerhuis et al. Geographic variability in outcomes within an international trial of glycoprotein IIb/IIIa inhibition in patients with acute coronary syndromes. Eur Heart J 2000; 21: 371–381.

Blair et al. Continental Differences in Clinical Characteristics, Management, and Outcomes in Patients Hospitalized With Worsening Heart Failure. J Am Coll Cardiol 2008;52:1640–8.

Chow et al. Assessing sensitivity and similarity in bridging studies. *Journal of Biopharmaceutical Statistics* 2002;12:385-400.

Fleiss. Analysis of Data from Multiclinic Trials. *Controlled Clinical Trials* 1986; 7:267-275.

The Hirulog and Early Reperfusion or Occlusion (HERO)-2 Trial Investigators. Thrombin-specific anticoagulation with bivalirudin versus heparin in patients receiving fibrinolytic therapy for acute myocardial infarction: the HERO-2 randomised trial. *Lancet* 2001; 358: 1855–63.

Higgins et al. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327;557-560

Kawai et al. An Approach to Rationalize Partitioning Sample Size Into Individual Regions in a Multiregional Trial. *Drug Information Journal 2008; 42. 139–147*.

Lachenbruch et al. Lot consistency as an equivalence problem. *Journal of Biopharmaceutical Statistics* 2004; 14:275–290.

Liu et al. Sample Size Requirements for Evaluation of Bridging Evidence. *Biometrical Journal* 2002;44:969–981.

Ministry of Health, Labour and Welfare of Japan, Basic Principles on Global Clinical Trials. September 28, 2007.

O'Shea et al. Statistical issues relating to international differences in clinical trials. *Am Heart J 2001;142:21-8.*

Pocock et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in Medicine* 2002; 21:2917–2930.

THE PURSUIT trial investigators. Inhibition of platelet glycoprotein IIb/IIIa with eptifibatide in patients with acute coronary syndromes. N Engl J Med 1998;339:436-43.

Quan et al. Sample Size Considerations for Japanese Patients in a Multi-Regional Trial Based on MHLW Guidance. *Pharmaceutical Statistics*, in press.

Rom et al. Testing for individual and population equivalence based on the proportion of similar response. *Stat Med* 1996; 15:1489-1505.

Shih J. Clinical trials for drug registrations in Asian-Pacific countries: proposal for a new paradigm from a statistical perspective. *Controlled Clinical Trials 2001; **22**:357–366.*

Stine, Heyse. Non-parametric estimates of overlap. *Stat Med* 2001;20:215–236.

Wedel et al. Challenges of subgroup analyses in multinational clinical trials: Experiences from the MERIT-HF trial. Am Heart J 2001;142:502-11.

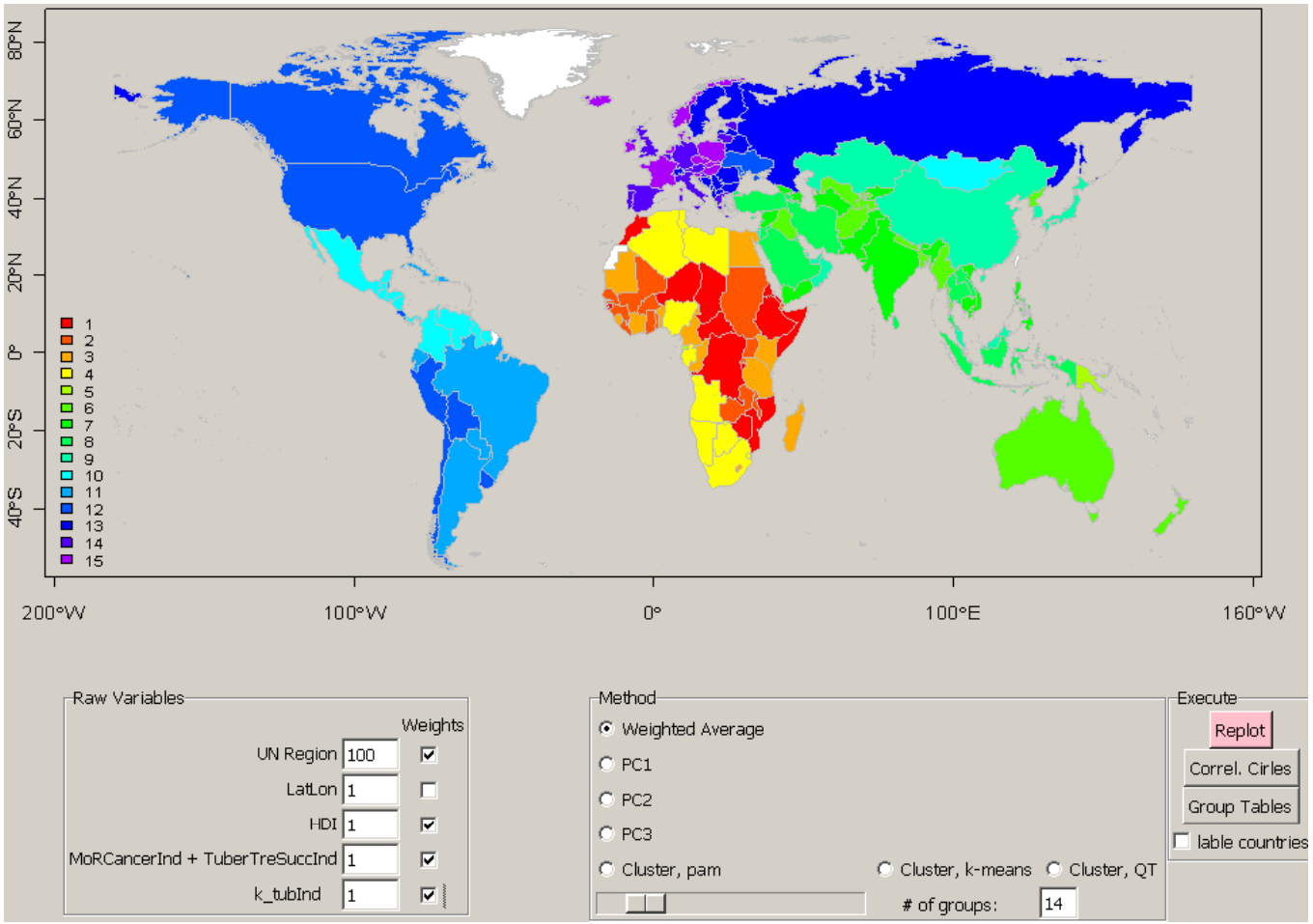
BACK-UP

PhRMA MRCT Workstream – Definition of Region – InnoCentiove Solution

- We can start from wide set (hundreds or thousands) parameters, characterizing countries. In our simple demo example we use GDP, Human Development Index, Mortality Rate of Cancer and Tuberculosis Treatment Success.
- We used two simplest ways to aggregate these variables to one measure:
  - Weighted sum or average = $\Sigma\ p.w_j$
  - Factor Analysis or Principal Component Analysis
    or
  - Clustering on full set on factors

PhRMA MRCT Workstream – Definition of Region – InnoCentiove Solution

R Program – Weighted Average, Factor Analysis, PCA, or cluster analysis



Equal weights, Weighted Average