Safeguarding public health

Multi-Regional Trials – A regulatory perspective

Andrew Thomson
Statistical Assessor, MHRA

Contents

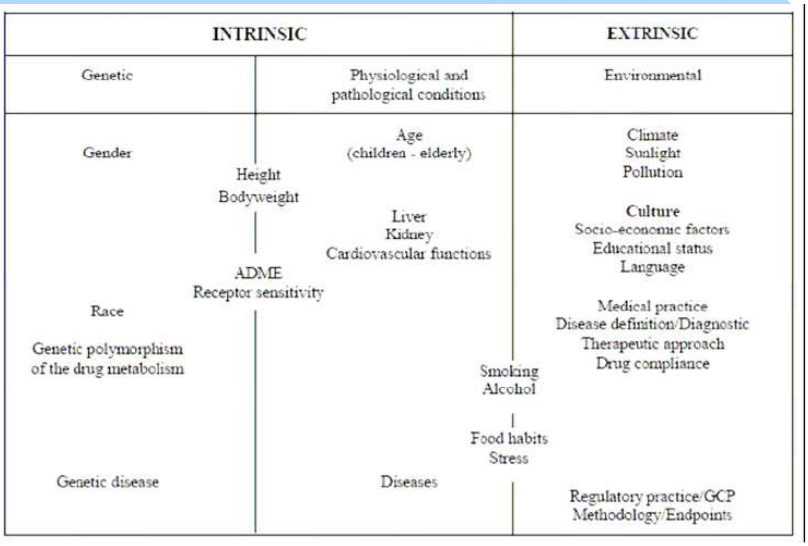
- Multi-regional pivotal trials of centrally authorised products
- Guidance documents – what messages for use of multi-regional trials in regulatory submissions
- Examples
- Planning and presenting confirmatory trials
- Multi-regional development programmes – an example

Guidance documents

- ICH E5 Ethnic Factors in the Acceptability of Foreign Clinical Data
  - "**All regions acknowledge the desirability of utilising foreign clinical data** that meet the regulatory standards and clinical trial practices acceptable to the region considering the application for registration. **However**…"
- CHMP REFLECTION PAPER ON THE EXTRAPOLATION OF RESULTS FROM CLINICAL STUDIES CONDUCTED OUTSIDE THE EU TO THE EU-POPULATION **(**EMEA/CHMP/EWP/692702/2008)
- Intrinsic (individual) and extrinsic (societal) factors identified.
- Indication-specific **"case-by-case"** consideration
- Guidance is **"cautiously inclusive"** but consider:
  - differences in regional standards (e.g. endpoints, control)
  - applicability / generalisibility of data
- No specific guidance on methodological aspects

Slide 3

Guidance documents



Slide 4

Applications with multi-regional (inc Europe) trials

MHRA

• Region can be thought of as defining a subgroup

• Common roles for subgroup analyses
  - Rescue a 'failed' trial
  - Investigate 'internal consistency'

• Different standards for assessment of each
  - Post hoc 'rescuing' rarely helpful (regulators appear to disbelieve subgroup findings)
  - Restrict some indications where internal consistency is not demonstrated (regulators appear to believe subgroup findings)

• **This is not a scientific divergence, but an exhibition of the precautionary principle**

Slide 5

Applications with multi-regional (inc Europe) trials

MHRA

• Strengths: what else do we have?
• Weakness: risk of misleading results / over-interpretation

• Case-by-case decision - critical is the strength and consistency of the data observed and the 'biological plausibility'
  - 'Differences in CV prevention by star sign' are implausible.
  - 'Differences in NSCLC tumour biology between Asian and Caucasian patients' are entirely 'plausible'

• Having multiple pivotal trials to estimate effects in the relevant subgroup is extremely helpful
  - Repetition is hard to ignore, even if biological plausibility not strong
  - Divergence often argues for a chance finding

• Subgroup analyses by region, where differences are plausible, are mandatory. Interpretation can be quantitative as well as qualitative.

Slide 6

Examples from Applications; Clopidogrel

- New indication sought for ST elevated MI (STEMI)
  - CLARITY; (6000 patients), 90 % Caucasian, surrogate endpoints
  - COMMIT; (46 000 patients), conducted in China, "hard" ("outcome") endpoints

	Clopidogrel plus aspirin	**Aspirin alone**	
Metoprolol	(i) **11500 patients** Active-clopidogrel plus aspirin + Active-metoprolol	(ii) **11500 patients** Placebo-clopidogrel plus aspirin + Active-metoprolol	Subtotal 1: 23 000 allocated **active-metoprolol**
No metoprolol	(iii) **11500 patients** Active-clopidogrel plus aspirin + Placebo-metoprolol	(iv) **11500 patients** Placebo-clopidogrel plus aspirin + Placebo-metoprolol	Subtotal 2: 23 000 allocated placebo-metoprolol
	Subtotal A: 23 000 allocated **active-clopidogrel** plus aspirin	Subtotal B: 23 000 allocated placebo-clopidogrel plus aspirin	

Figure (9.1) 1 - Factorial design among 46 000 patients

Slide 7

Examples from Applications; Clopidogrel

The COMMIT study; results

Event	No. (%) With Event: Clopidogrel 75 mg* (N = 22961)	No. (%) With Event: Placebo* (N = 22891)	Odds Ratio (95% CI)	Absolute Benefit /1000 (SE)	Two-sided p-value[a]
Composite endpoint: Death, re-MI, or Stroke[b]	2121 (9.2%)	2310 (10.1%)	0.91 (0.86, 0.97)	8.5 (2.8)	0.002
Death	1726 (7.5%)	1845 (8.1%)	0.93 (0.87, 0.99)	5.4 (2.5)	0.029
Nonfatal re-MI[c]	270 (1.2%)	330 (1.4%)	0.81 (0.69, 0.95)	2.7 (1.1)	0.011
Nonfatal stroke[c]	127 (0.6%)	142 (0.6%)	0.89 (0.70, 1.13)	0.7 (0.7)	0.333

* All treated patients received daily ASA (162 mg).

[a] Based on log-rank test.

[b] The difference between the composite endpoint and the sum of death + nonfatal re-MI + nonfatal stroke indicates that 9 patients (2 clopidogrel and 7 placebo) suffered both a nonfatal stroke and a nonfatal re-MI.

[c] Nonfatal re-MI and nonfatal stroke exclude patients who died (of any cause).

SE = standard error; re-MI = reinfarction.

Slide 8

Examples from Applications; Clopidogrel
COMMIT results

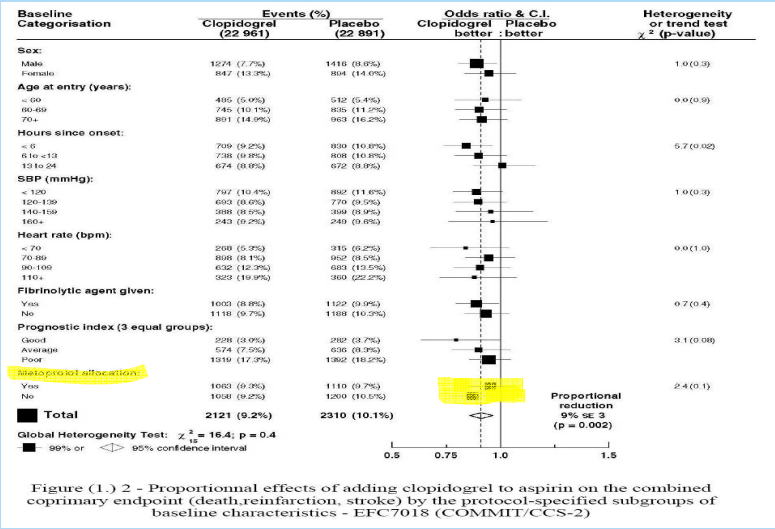


Figure (1.) 2 - Proportionnal effects of adding clopidogrel to aspirin on the combined coprimary endpoint (death,reinfarction, stroke) by the protocol-specified subgroups of baseline characteristics - EFC7018 (COMMIT/CCS-2)

Slide 9

Examples from Applications; Clopidogrel

- Concerns regarding relevance of the clinical setting to EU clinical practice;
  - Background therapy; most European patients get beta-blockers
  - Results may not be relevant to European population

- CHMP considered COMMIT trial as supportive rather than pivotal

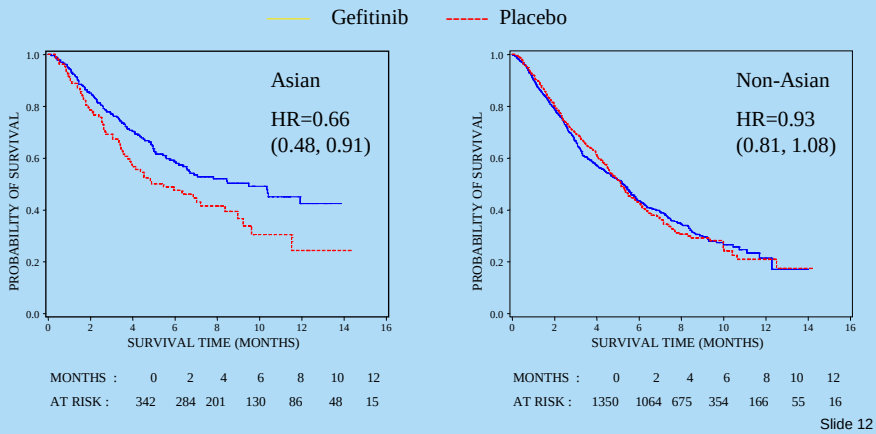
Slide 10

Examples from Applications; Gefitinib

- Non-small cell lung cancer refractory to other therapies
- Two pivotal trials in initial application;
  - One mainly in Caucasians, one in a mixed Asian and Caucasian population
  - Data suggested a higher response in Asian patients
- Following the above one large study, ISEL, was conducted including 75% Caucasians

Slide 11

Examples from Applications;Gefitinib
Survival Data by Ethnicity



Slide 12

Examples from Applications;Gefitinib

- Based on these data:
  - survival effect only demonstrated in Asian patients
  - no obvious explanation for the lack of survival effect in Caucasian patients
  - hypothesised to be because of tumor genetics?
  - Gefitinib approved in Japan but not in the EU

- Further studies
  - Gefitinib effective predominately in mutation positive tumors
  - Asian patients; higher prevalence of mutations (40 versus 10%)
  - Caucasians; satisfactory effect in patients with mutations

Slide 13

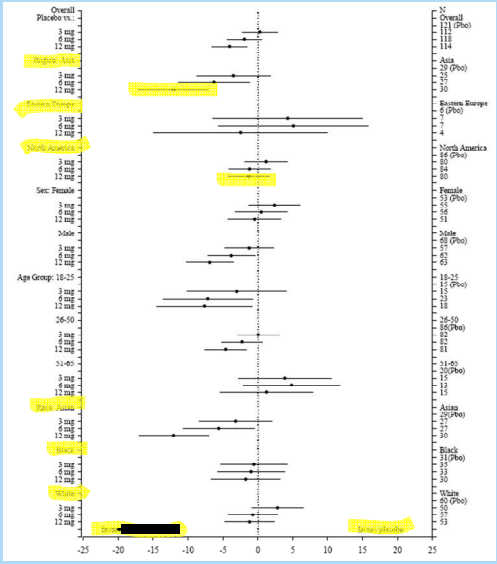
Examples from Applications; Anonymised Antipsychotic

- Treatment of acute manic episodes associated with bipolar disorder
  - One fixed dose study and one flexible dose study;
  - Majority patients from US (71% and 54% respectively)

- Significant effects in favour of test drug compared to placebo
  - No effect in US patients
  - Fixed dose study; large effect in Asian, primarily Indian patients (15% of the total study population)
  - Flexible dose study; large effect in Russian and Ukrainian patients (29%)

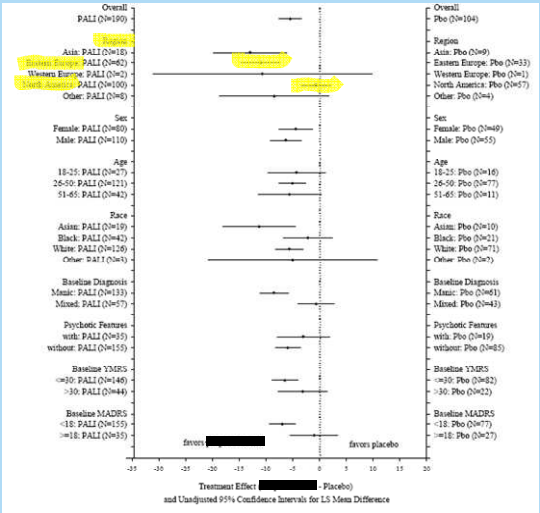
Slide 14

Examples from Applications; Anonymised Antipsychotic
Results fixed dose study



Slide 15

Examples from Applications; Anonymised Antipsychotic
Results flexible dose study



Slide 16

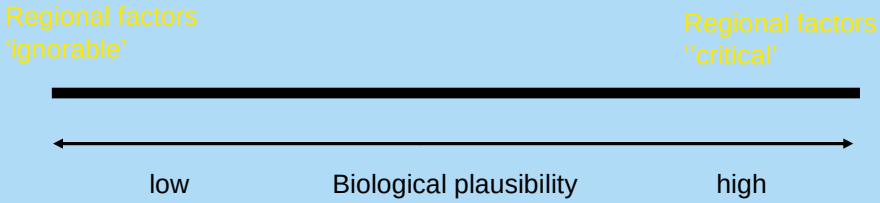
Examples from Applications; Anonymised Antipsychotic

- CHMP conclusion:
  - Questionable whether positive findings are representative for the major part of European patients suffering from an acute episode of mania because of the differences in race and/or medical and social environment
  - Application withdrawn

Slide 17

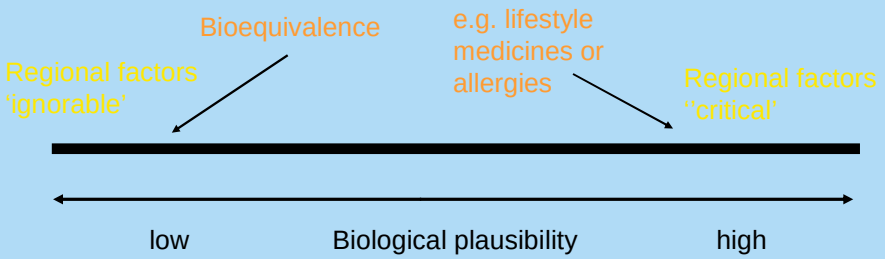
Planning confirmatory trials

- For what reason might heterogeneous effects of treatment by region be hypothesised? To what degree are these plausible?
- Discuss in protocol (not influenced by data). Consider to stratify randomisation.
- Seek Scientific Advice if you are concerned about regulatory interpretation



Slide 18

Planning and presenting confirmatory trials



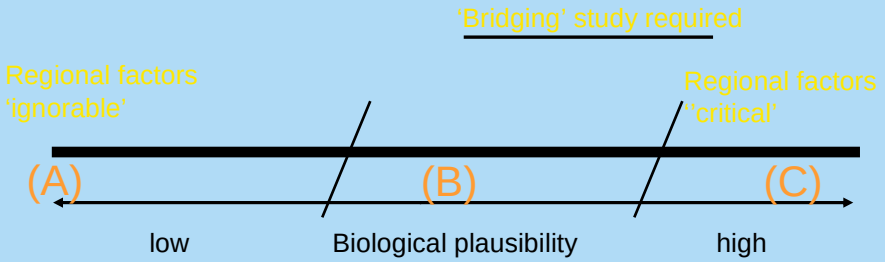
Slide 19

Planning and presenting confirmatory trials

A – no specific statistical considerations for planning

B – where feasible, plan to demonstrate that effects are consistent trusting to luck or using, for example, α = 20% and / or bridging strategy

C – plan for stand-alone evidence



Slide 20

Applications with multi-regional development programmes

- With no European data in a trial, the issue is less clear
- What countries can we borrow strength from?
- “plan to demonstrate that effects are consistent trusting to luck or using, for example, α = 20%” no longer applicable. Makes bridging much more problematic.
- Intrinsic and extrinsic factors become even more important
- Active comparator appropriate?
- Stand alone European data may be required

Slide 21

Applications with multi-regional development programmes.

• Oral contraceptives - The study size requirements and pregnancy reporting of the Notes for Guidance on clinical investigation of steroid contraceptives in women regarding efficacy for a new contraceptive method include:

− The calculation of efficacy is based on the Pearl Index
− The difference between the point estimate and the upper limit of the 95% confidence interval does not exceed 1
− At least 400 women have completed one year of treatment

• Expected that the Pearl Index point estimate will be below 1

Slide 22

Example - Qlaira

- Qlaira – estradiol valerate + dienogest (DCP - NL as RMS)
- Data from both the EU and the US population
- Well known that the Pearl Index is higher in the US
- Qlaira is not unique in this respect – applies to all routes of administration of hormonal contraceptives.
  - Genetic?
  - Education?
  - Weight?
  - Motivation?

Slide 23

Example - Qlaira

The unadjusted Pearl Index (PIU) based on pooled data across 2 studies conducted in the **EU** for women aged 18-35 years is **0.87** (upper limit 95% CI: **1.52**)

Taking the **US** study separately, the unadjusted Pearl Index (PIU) for women aged 18-35 years is **1.45** (upper limit 95% CI: **3.16**)

Combined with the European studies the unadjusted Pearl Index (PIU) for women aged 18-35 years is **1.01** (upper limit 95% CI **1.59**)

Slide 24

Example - Nuvaring

Nuvaring – vaginal insert

2 studies – one EU (mainly Scandinavia) one US

Results similar but combined less clear: (MEBs PAR)

Group	Trial	N	Total extent of exposure		In-treatment pregnancies	Pearl-Index estimate	95% CI
			Number of 28-day cycles	Woman years			
ITT[a]	068003	1,177	11,188.1	857.7	15	1.749	0.979 - 2.885
	34219	1,145	12,109.4	928.3	6	0.646	0.237 - 1.407
	Combined	2,322	23,297.6	1,786.0	21	1.176	0.728 - 1.797

Slide 25

Example - Nuvaring

Why the difference between regions?

Compliance – 80% US versus 91% EU

Temporary removals occurred more often in US (25% versus 10% in EU)

- No difference for short term removal (0 – 4 hours)
- Bigger difference for >4 hours

Differing advice on use of back-up contraception in trials.

Slide 26

Example - Nuvaring

US and the European study populations were rather similar, except for the number of gravid women and black race.

Extensive data on body weight provided indicated no clinically relevant correlation between the plasma levels of ENG and EE and the body weight of a subject using NuvaRing.

Race a predictive factor for lack of compliance – confounding?

Which data is most relevant to the general European population, given the explanations provided by the Company?

Slide 27

Example - Anya / Lybrel

Anya / Lybrel - Licensed in US – From the FDA website:

“The primary efficacy and safety study (313-NA) was a one-year open-label clinical trial that treated 2,134 subjects in North America.

The resulting total Pearl Index was 2.38 (95% CI: 1.51, 3.57)

In a second supportive study conducted in Europe 641 subjects were randomized to LYBREL (n=323) or the cyclic comparator (n=318) There was one pregnancy in the LYBREL group that occurred within 14 days following the last dose. There were three pregnancies in the cyclic comparator group.

Slide 28

Example - Anya / Lybrel

3 pregnancies in comparator group, 0 in Anya group

Note that FDA considers US study as primary

Certainly where the vast majority of the data to support efficacy was demonstrated

European data much more limited, despite Pearl Index poitn estimate of 0

Licensed in one EU country – MRP followed, ended up with withdrawal of product entirely from European market

Slide 29

Example – Contraceptives

- Products considered on a case-by-case basis
- Even if identical indications
- May not be sufficient to just do a small European study
- Principle extends across treatments and indications.
  - Good stand-alone data may be required
  - May be difficult to ignore the other data that is 'supportive', without a robust explanation of the differences
- In this respect, no different to multi-regional *trials*

Slide 30

Conclusions and Questions

- Evidence coming from multiple regions can be acceptable. Consider regional standards as well as generalisability of trial findings
- Differences in intrinsic/extrinsic factors and study conditions exist both within and between regions. It is important to identify, control (where possible) and investigate relevant factors
- Failure to prospectively consider potential regional differences, and to plan for evidence to address potential heterogeneity, represents a risk to the drug developer.

Any questions?

Acknowledgements to Rob Hemmings for some of the slides, who in turn acknowledges Kristina Dunder, Barbara van Zweiten Boot, Tomas Salmonson for ideas, case studies and slides.