Multi-regional Regulatory Guidance on Multiple Endpoints in Clinical Trials

Yong Wang, PhD

Aug 30 2018

PAREXEL®
YOUR JOURNEY. OUR MISSION.™

Outline

- Introduction
- Define claim and how far should “claim” be made
- Composite endpoints
- Subgroup analysis
- Dose-finding analysis
- Common statistical methods
- Challenges

European Medicines Agency

September 1998
CPMP/ICH/363/96

ICH Topic E 9
Statistical Principles for Clinical Trials

Multiple Endpoints in Clinical Trials

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER) Scott Goldie at 301-796-2055 or (CBER) Office of Communication, Outreach, and Development, 800-835-4709 or 240-402-8010.

· CCTS 专家共识 ·

临床试验中多重性问题的统计学考虑

CCTS 工作组 王彤 易东 执笔

临床试验根据研究目的不同可分为“探索性临床试验”和“确证性临床试验”，而临床研究结论通常需要根据确证性临床试验的统计推断结果得到。如果某一确证性临床试验需要对多个检验假设做出统计学推

所示。

表 1 多重检验的结果

原假设	拒绝	拒绝	合计
真	U	V	m_0

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

15 December 2016
EMA/CHMP/44762/2017
Committee for Human Medicinal Products (CHMP)

Guideline on multiplicity issues in clinical trials

Draft

Draft agreed by Biostatistics Working Party (BSWP)	November 2016

Defining "claim"

- Ensure that all 'claims' have *overall* Type I error rate control
  - Study-wised type I error
  - In "strong" sense, i.e. there is control on the probability to reject at least one true null hypothesis, regardless which subset of null hypotheses happens to be true. (EMA 2002)
  - FDA's concern for controlling the Type I error probability is to minimize the chances of a false favorable conclusion for any of the primary or secondary endpoints, regardless of which and how many endpoints in the study have no effect (called strong control of the Type I error probability).

Defining “claim”

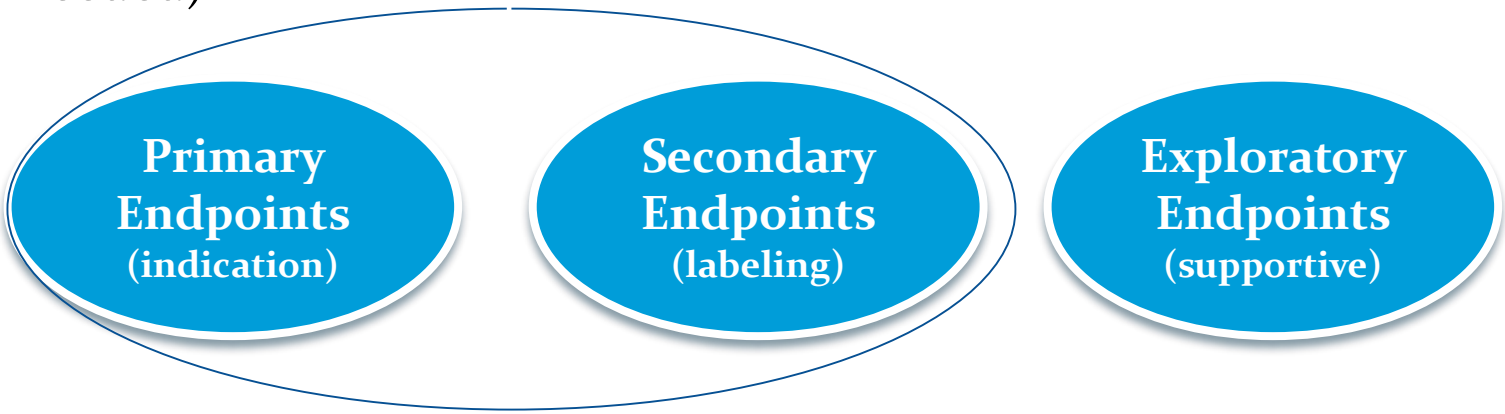
EMA:

- The reader *should not* directly relate use of the word claim with the possibility to make statements or present data in the Summary of Product Characteristics, which is governed by a separate regulatory guidance document.
- “claim” is used as (1) shorthand for a confirmatory conclusion which is then prioritised in a clinical study report, clinical overview or clinical summary, and (2) is used as a primary basis for asserting that efficacy or safety has been established.

How far should "claims" be made

To how far, a hypothesis is needed

- (FDA, 2017) *All primary and secondary* efficacy endpoints need to be tested and type-I error needs to be controlled. Other endpoints need to be put into the exploratory endpoint (only descriptive stat is needed)



- Recommend the list of secondary endpoints be short

How far should “claims” be made

- **FDA**
  - The focus: control of the Type I error rate for the planned primary and secondary endpoints.
  - Type I error should be controlled WITHIN and BETWEEN primary and secondary endpoint families
  - *Presenting p-values from descriptive analyses* (that is, from analyses that were not prespecified and for which appropriate multiplicity adjustments were not applied) *is inappropriate*

How far should "claims" be made

- (EMA, 2002) **Secondary endpoint**: there has been no common consensus about the role and the weight of secondary endpoints in clinical trials
- (EMA 2017) **Secondary endpoints:**
  - *Expressing supportive evidence* ---
    - No claims are intended; confidence intervals and statistical tests are of descriptive nature
  - *May become the basis for additional claims* ---
    - Significant effects in these endpoints can be considered for an additional claim
      - 1. only after the primary objective of the clinical trial has been achieved, and
      - 2. if they were part of the confirmatory strategy.

How far should “claims” be made

- (continue) general procedure:
  - hierarchically primary -> secondary
  - more complex methods exist to control type I error over both primary and secondary endpoints. Regulatory dialogue is recommended to assure that the outcome of the procedure can be interpreted in clinical terms.

Composite endpoint

- (FDA and EMA) Each individual component needs to be evaluated, however, *NO need for multiplicity adjustment* after significance of primary endpoint (composite endpoint) is achieved.

- Treatment should be expected to affect ALL components in a similar way.
  - For any component that is included in the composite, it is usually appropriate that any additional component reflecting a worse clinical event is also included. (e.g. if hospitalization is a component, then more adverse clinical outcomes, such as MI, stroke, death, should be included as well.

- (EMA) Non-inferiority / equivalence studies --- composite endpoints pose particular issues
  - Adding a component that foreseeably is insensitive to treatment effects tends to decrease sensitivity of the comparison

Composite endpoint

- (FDA) Example of presenting a composite endpoint
  - RENAAL study (losartan to delay development of diabetic nephropathy)

Table 1. Decomposition of Endpoint Events in RENAAL*

Endpoint	Losartan (N=751)	Placebo (N=762)	Hazard ratio± (95% CI)	p-value
Primary endpoint				
Doubling of serum creatinine, ESRD, or death	327	359	0.84 (0.72, 0.97)	0.022
Decomposition of the primary endpoint				
Doubling of serum Creatinine	162	198	0.75	
ESRD	64	65	0.93	
Death	101	96	0.98	
Any occurrence of individual components				
Doubling of serum Creatinine	162	198	0.75 (0.61, 0.92)	
ESRD	147	194	0.71 (0.57, 0.89)	
Death	158	155	1.02 (0.81, 1.27)	

*Excerpted from FDA/CDER/DBI Statistical Review at (http://www.accessdata.fda.gov/drugsatfda_docs/nda/2002/20-386s028_Cozaar.cfm).

ESRD = end-stage renal disease; ±Hazard ratio from Cox proportional hazards time-to-event analysis.

Safety evaluation

- (EMA)Multiplicity in safety variables
  - When a safety variable is *part of the confirmatory strategy* of a study and thus has a role in the approval or labelling claims, it should not be treated differently from the primary efficacy endpoints
  - A non-significant difference between treatments will not allow for a conclusion on the absence of a difference in safety
  - Due to the precautionary principle in safety evaluations, reducing the *rate of false negative conclusions* on harm is usually more important than controlling the number of false positive conclusions and rigorous multiplicity adjustments could mask relevant safety signals

Safety evaluation

- (FDA) The multiplicity problem is an issue in safety evaluations of controlled trials
  - no prior hypotheses
  - many plausible analyses
  - numerous safety findings that would be of concern
  - interest in both individual large studies and pooled study results.
- There is no easy remedy for these issues
- it is more credible that there is a causal relationship between an observed adverse event and the drug

… …

The multiplicity problems for these types of safety analyses are outside the scope of this guidance.

Subgroups Analysis

(EMA) Reliable conclusions from subgroup analyses generally require pre-specification and appropriate statistical analysis strategies

A specific claim of a beneficial effect in a particular subgroup requires pre-specification of the corresponding null hypothesis

Dose-finding Analysis

- (EMA) Does-finding study: usually statistical inference should focus on estimation rather than testing.

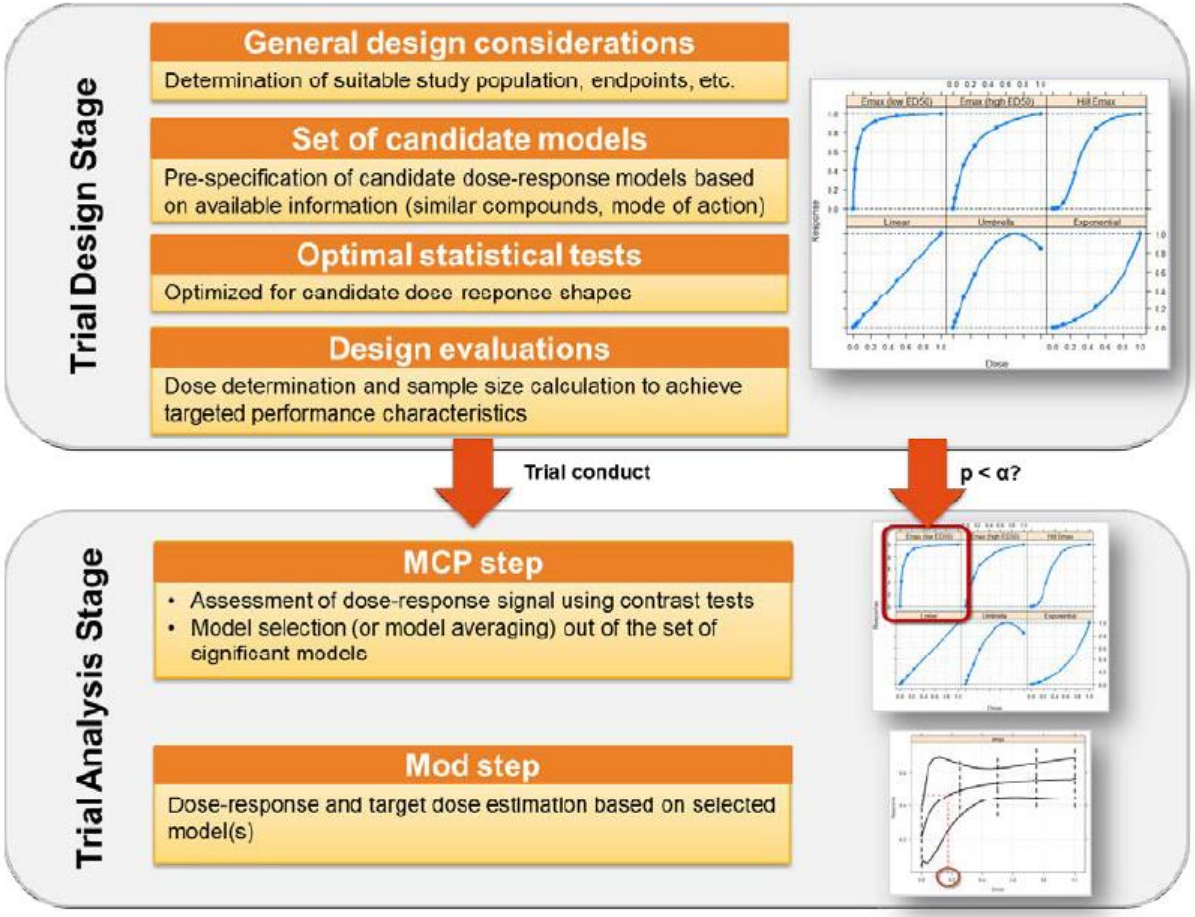
- the multiplicity adjustment of the different comparisons between groups in order to control the study-wise type I error may *not* be required in a Phase II trial

- if study aim is to identify one dose in specific patient population, then type-I error control is mandatory

Dose-finding Analysis

- (EMA 2014) *Qualification Opinion of **MCP-Mod** as an efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty*

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH



Common Statistical Methods

- FDA and CFDA proposed numerous statistical methods: for example,
  - *Bonferroni Method*
  - *Holm Procedure*
  - *Hochberg Procedure*
  - *Gatekeeping Testing Strategies*
  - *Resampling-Based, Multiple-Testing Procedures*
  - 多剂量组与对照组相比 (Dunnett method) and 多个剂量组相比，无安慰剂和阳性对照 (Shaffer method)

Challenges

- Handle situations where different regulatory bodies requests different primary endpoints
- Subgroup analysis
- Are methods that pass α from secondary endpoints back to primary endpoints permissible?
- How much justification is needed for some commonly used procedure which relies on other assumptions (e.g. Hochberg test)

THANK YOU

