Reflection paper on statistical methodology for the comparative assessment of quality attributes in drug development

J-M. Cardot

Email: J-michel.cardot@uca.fr

UFR DE PHARMACIE
Université Clermont Auvergne

Introduction

dates

- EMA/CHMP/138502/2017 23 March 2017
- Draft agreed by Biostatistics Working Party February 2017
- Adopted by CHMP for release for consultation 23 March 2017
- Start of public consultation 01 April 2017
- End of consultation (deadline for comments) 31 March 2018

Structure

- 7 parts
  - 1. Introduction
  - 2. Legal basis and relevant guidelines
  - 3. Definitions and delineations
  - 4. Settings where the comparison on the quality level is of particular relevance in regulatory decision-making
  - 5. Approaching the comparison task from the statistical perspective and associated obstacles
  - 6. Reflections of issues raised, implications for planning and assessment
  - 7. Appendix

Scope-area in brief

- Realistic requirements to demonstrate ‘similarity on the quality level’ during
  - drug development,
  - drug lifecycle,
  - decision making processes potentially leading to marketing authorisation
- Area
  - pre- and post-manufacturing change,
  - biosimilar developments
  - generics development
- Methodological aspects in relation to statistical data-comparison
  - statistical perspective comparison objectives,
  - sampling strategies,
  - sources of variability,
  - options for statistical inference and acceptance ranges.
- Connect to other regulatory guidance comparing quality attributes and/or improving methodology when lacking

To summarize with Key words

- Quality
  - Quality Attribute: QA
  - Process control methodology and system
  - Deviation from expected quality, similarity of quality
  - Improvement of quality, link with consistency
- Statistics
  - Data distribution,
  - Similarity of variances, of central parameters: equality, non inferiority, difference
  - Sampling
  - Limits setting

Based on / Linked to

- ICH Q5E (Biologicals), Guideline on similar biologicals (CHMP 437/04/rev1 and EMEA/CHMP/BWP/49348/2005),
- Q8-11 and later on Q12
- Guideline on BE (CPMP/EWP/QWP/1401/98)
- Guideline on MR (EMA/CHMP/EWP/280/96)
- Guideline locally applied, locally acting products (CPMP/EWP/239/95)

This presentation

- Focus on NCE and generics
- Not focused on biologics

Brief description of the problem (if any)

Scope

- The comparison of a particular drug product in versions pre- and post-manufacturing change (EU-SUPAC ?)
- *The comparison of a candidate biosimilar product to a reference medicinal product*
- The comparison of a candidate generic product to the reference medicinal product (Development)

=> support the assertion that the quality profile of two (versions of a) drug products can be considered similar

Why

- Similarity
  - Classical “inferential” statistical methods aim show difference and not similarity
  - The lack of significant differences alone does not imply similarity => function of power (1-β), N etc…
  - Limits based on ???? (ex content based on pharmacopeia!).
- Extrapolation: Limited information from sample data => not a lot of batches, values, often sequential (first new batches vs last old batches, etc…).

Why

- Tools used to measure
  - Quantitative: precision, accuracy, sensitivity, reproducibility, reproducibility, etc…
  - Qualitative white to off white…
- Limits used
- Comparison driven by non statistical tools case by case

=> Is the set of QA known and can I measure them accurately can I conclude with a priori justified test and limits

Questions

- Questions: non inferiority, equivalence, difference?
- If equivalence or non inferiority/superiority how to set limits of acceptance
- Number of units to insure test validity and to be able to conclude/extrapolate

Aim

- Compare quality levels
- Find a common approach that
  - Sound statistically correct
  - Could be used in practice
  - Protect Patient
  - Allows continuous improvement of quality (??)
  - Has a scientific background
- And is still manageable by users!

Example actual

Position

- Compare quality of two product Test and Ref

⇒ensure the quality, safety and efficacy of drug product

- Insure that QA are
  - Similar
  - Improved (for example impurities)

=> no negative impact on safety and efficacy (positive impact possible)

- Problem
  - Number of batches
  - Sampling
  - Unit used
  - Type of essays and sensitivity

Actual approaches

- One or 2 batches, not randomly sampled
- Tolerance interval (TI), x-sigma (example: 2 x sd) min-max range (example mass of tablets) => no clear conclusion
- Limits based on 0.8000-1.2500 ???
- Limits based on texts, usage, etc… and not always on science

Example proposed

Hypothesis and problems

- Hypothesis
  - non inferiority,
  - Superiority
  - Equivalence
  - Difference
- If superiority, equivalence or non inferiority how to set limits of acceptance
- Number of units to insure test validity and to be able to conclude/extrapolate

Sampling

- Sample must be representative => random sampling but hardly feasible: limited number of batches (consecutive?), large number must keep samples, aging/shelf life influence
  - Consistent manufacturing process
  - Known source of variability
  - Sampling/samples must bring information
- Non random => representative ? If question how to extrapolate to all further batches
- Pseudo random => set up strategy based on pre defined assumptions of representativeness

Criteria, acceptance range

- Criteria/Acceptance range must be defined a priori and not derived from data under interest but from previous set of data (a priori knowhow)
- Acceptance limits in the protocol before the study
- Function also of the distribution
- Function of the possible clinical outcome or good pharmaceutical quality the stricter of the two.
- Sometime arbitrary

Success criteria

- Often more than one QA => more than one statistics
  - Qualitative
  - Quantitative
- Set up an a priori success concept binding all criteria
- No post hoc justification
- Risk false positive
- Risk of alpha inflation
- Post hoc power calculation (more than sample size calculation …!)

Quality Attribute

- Unknown distribution(s): test and ref
- Qualitative or quantitative
- Quantitative
  - Central position: mean (?)
  - Dispersion: variance (?)
- Need to know distribution characteristic before planning tests
- One sided (ex: reduction of impurities) or two sided (ex: “absence” of difference in content) => needed before test

Proposed example transfer/variation

- QA after chances “non inferior” to QA before change
- Representative sample of units
  - Larger set of initial (pre change) units (batches)
  - Post manufacturing could be limited and consecutive
  - Could help to see consistency post change=> must be OK
- Batch number (3 cited but not justified)
- Statistical model to identify source of variation of both production (formulation etc…) and assay => know the within et between sources of variabilities
- Justification of limits/specifications needed

Proposed example dissolution

- Batch to batch consistency
- Justification of waivers

=> Inferential idea, similarity in dissolution from tablet sample could be extrapolated to population(s) even after scale up

- Single unit dissolutions (n=?? 12??) but no mention of sampling points to be used
- No mention in case of different variability between test and ref and sources of variability
- No mention on the ad equation of the dissolution method for both test and ref

Proposed: dissolution

- F2
  - Use mean, and based on average difference
  - Insensitive to time interval
  - No shape comparison
- When F2 not possible other distance based method used
  - Raw data
  - After modeling ….
- Always based on central parameters … mean value

Proposed: dissolution

- F2 no alpha (and of course no beta) risk associated … exept after bootstrapping
- F2: not possible to make a simple CI (except if bootstrapping)
- F2 acceptance based on a mean almost 10% difference
- Alternative to F2
  - Limits +/- 10% … of what (ref ?)
  - Limits +/- 10% of biobatch but why
  - What is the in vivo outcome of +/-10%
- How to set alpha risk … problem of multiple comparison R1 vs T1 R1 vs T2 R1 vs T3, R2 vs T1, etc…

Discussion

ICH Q12

- Does this reflection paper prepare ICH Q12?
- Yes as in this case it will be a paper based dossier post modification in some cases

⇒All QA known

⇒All under Quality (ICH Q8-11)

⇒Close to SUPAC

Is continuous improvement possible

- In pharmacy that means that product is not of constant quality
- Could increase robustness but must insure similar clinical outcome in safety and efficacy

Justification / Phamacopeia

- How to deal with pharmacopeia … is this “book” obsolete
- Could not base any more on it for limits
- => CoA … limits may be next step link with this guideline?

Items

Check list or decision pathway?

- General description of comparison setting/comparison objectives
- Given the QAs of interest, categorisation of QAs regarding scale of measurement (binary to continuous)
- For each QA, decision upon the characteristic/parameter of interest by which 906 underlying data distributions will be compared (e.g. mean, variance, etc.)
- Translation to statistical objectives, e.g. deciding upon one- or two-sided comparison approach per QA
- Identification of the unit of observation; at the same time exploration of potential sources of variability in QAs' data to be
- Consideration for which potential sources of variability the data analyses can be controlled for Sampling strategy
- Definition of metric/method to describe difference/distance between the chosen parameters (e.g. difference in means, ratio of means, etc.)
- Evaluation whether the so chosen setup for QA data comparison would allow for inferential statistical approach
- Pre-specification of an acceptance range for the analysis of each QA separately (e.g. equivalence margin, non-inferiority margin)
- Consideration regarding the risk for a false positive conclusion on similarity (equivalence/non-inferiority) based on the similarity decision criteria defined

Thank you

Questions ?

No => perfect ! ☺

Yes => I am ready to answer! ☹ … ☺