Is Similarity Different?

Recent developments from the regulatory perspective

Stephan Lehr
Austrian Medicines & Medical Devices Agency

The regulatory perspective reflects a personal opinion, remarks do not necessarily reflect the view of the agency

Outline

Have a guess!

"Due to the high PK variability and a rather flat exposure-response relationship, a comparability range of up to 66-150% is considered acceptable but the point estimate will also be taken into account when assessing PK similarity."

"The choice of estimands for studies with objectives to demonstrate non-inferiority or equivalence requires careful reflection."

"Bayesian methods are a further source of 'adding assumptions' to data. […] Such methods may be advantageous […] although introducing prior beliefs is often a concern in drug regulation."

"There are no special methods for designing, carrying out or analysing clinical trials in small populations."

AGENDA

Is similarity different?

- Revision of „products containing recombinant granulocyte-colony stimulating factor" Guideline
- Estimand in clinical equivalence trials
- Bayesian approaches to demonstrate biosimilarity
- Development in rare diseases

Revision of product specific guideline

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

26 July 2018
EMEA/CHMP/BMWP/31329/2005 Rev 1
Committee for Medicinal Product for Human Use (CHMP)

Guideline on similar biological medicinal products containing recombinant granulocyte-colony stimulating factor (rG-CSF)

Draft

Draft agreed by BMWP	July 2018
Adopted by CHMP for release for consultation	26 July 2018
Start of public consultation	15 August 2018
End of consultation (deadline for comments)	15 February 2019

Revision of rG-CSF guideline

Specifics for pegylated rG-CSF

165 In principle, cross-over or parallel-group single-dose PK studies are acceptable. The intra-subject
166 variability of pegylated rG-CSF PK is considerably lower than the inter-subject variability. Hence, cross-
167 over studies decrease the notably high PK variability of pegylated rG-CSF but require a long wash-out
168 phase (at least 6 weeks) to avoid relevant carry-over of pharmacological effects. Studies with a parallel
169 group design on the other hand will require a higher number of study subjects to account for inter-
170 subject variability.

- Parallel group designs: no impact of residual PK/PD effects, immunogenicity clearly attributed to drug administered
- Cross-over analysis and interpretation, e.g. treatment-by-period interaction

Revision of rG-CSF guideline

Specifics for pegylated rG-CSF

174 Injection site and injection technique should be standardised to decrease variability. Other factors that
175 may affect drug exposure are bodyweight and potentially anti-PEG antibodies. If pre-planned, a
176 subgroup analysis of subjects without pre-existing or treatment-emergent anti-PEG antibodies may be
177 acceptable for proof of similar PK profiles. However, a large difference in antibody development may
178 question biosimilarity.

- Treatment emergent anti-PEG antibodies are post-randomisation events, and a subgroup analysis may be difficult to interpret

Revision of rG-CSF guideline

Specifics for pegylated rG-CSF

185 Due to the high PK variability and a rather flat exposure-response relationship, a comparability range
186 of up to 66-150% is considered acceptable but the point estimate will also be taken into account when
187 assessing PK similarity.

- PK variability sufficiently understood?
- Dose-dependent increase of safety signals observed in Neulasta
- Poor association between PK- and PD-differences within subjects
- Limits sufficiently justified in light of possibility for a tailored clinical program (no comparative efficacy trial)
- Q&A document: "… a point estimate or substantive part of the confidence interval lying towards the extremes of the acceptance criteria will require further discussion."

Estimands in Equivalence Trials

ICH E9 (1998)

- „… in an equivalence or non-inferiority trial use of the full analysis set is generally not conservative and its role should be considered very carefully. “
- “The equivalence (or non-inferiority) trial is not conservative in nature, so that many flaws in the design or conduct of the trial will tend to bias the results towards a conclusion of equivalence.“
- PPS: “… test of the hypothesis and estimate of the treatment effect may or may not be conservative depending on the trial; the bias, which may be severe, arises from the fact that adherence to the study protocol may be related to treatment and outcome

Estimands in Equivalence Trials

PtC switching between superiority and non-inferiority (2000)

- "In a non-inferiority trial the full analysis set and the PP analysis set have equal importance and their use should lead to similar conclusions for a robust interpretation."

Estimands in Equivalence Trials

ICH E9 (R1) Addendum, DRAFT (2019)

- **Intercurrent Events (IEs)**

  Events that occur after treatment initiation and either affect interpretation of the variable or preclude its observation.

- **Estimand**

  Target of estimation to address the scientific question of interest posed by the trial objective. Attributes of an estimand:

  - population
  - variable (or endpoint)
  - treatment condition(s)
  - specification of how IEs are reflected in scientific question
  - population-level summary for variable

Estimands in Equivalence Trials

ICH E9 (R1) Addendum, DRAFT (2019)

- Estimand addressing intercurrent events (non-adherence, other medication, death) using the treatment policy strategy present same issues as FAS analysis: responses in both arm will appear more similar following treatment discontinuation or use of another medication
- => construct estimand that directly addresses those intercurrent events
- => construct estimand that prioritises sensitivity to detect differences between treatments, if they exist

Estimands in Equivalence Trials

ICH E9 (R1) Addendum, DRAFT (2019)

- Selecting strategies, it might be important to distinguish between trials designed to detect whether differences exist between formulations (e.g. comparison of a biosimilar to a reference treatment) and trials where a non-inferiority or equivalence hypothesis is used in order to establish and quantify evidence of efficacy.
- How to address the potential anti-conservative nature?
  - Co-primary estimands in accordance to FAS and PPS?
  - "Sensitive estimand with supplementary "clinically relevant" analysis or clinically relevant estimand with supplementary "sensitive" analysis"?
  - Agreement on strategies (to address IEs) and assumptions in estimators for estimands?

Estimands in Equivalence Trials

ICH E9 (R1) Addendum, DRAFT (2019)

- PPS analysis does estimate effect in adherent subjects (it may not compare similar subjects in different arms)
  - What's the role of PPS analysis, what does it estimate?
  - Some protocol violations and deviations might be addressed as intercurrent events.
  - Where a majority of intercurrent events are reflected in the construction of the estimands, the number of remaining protocol violations and deviations will be low and analysis of the PPS might not add additional insights.
- Having specified the population and the strategy to address intercurrent events, what's the role for analysis sets?

Bayesian Methods

Incorporating prior information into phase III trial

- Idea:
  Assessment of strength of prior evidence (analytical, PK/PD similarity) is borrowed for analysis of clinical equivalence study
- Hope:
  More efficient statistical analysis of efficacy

- Examples of methods proposed
  - Weiss et al (2018) „Bayesian methods for analysis of biosimilar phase III trials", StatMed, 37(20):2938-2953
  - Zeng et al (2017) „Improving the power to establish clinical similarity in a Phase 3 efficacy trial by incorporating prior evidence of analytical and pharmacokinetic similarity", JBiopharmStat 2018;28(2):320-332

Bayesian Methods

Weiss et al (2018)

- If analytical and PK/PD similarity has been shown:
  - model π_B having prior mean π_R with variance σ_B^2 to be specified
  - (π_B , π_R response rates on biosimiliar and reference product, respectively)
- π_R determined by meta-analysis, and set $\sigma_B = S_{min} \cdot W_c \cdot W_I$ (S_{min}=0.05)

			Analytical testing				
			Weak support	similar	Highly similar	Fingerprint similar	
			X	2	1.5	1	**W_c**
Phase I	Weak support	X	X	X	X	X	
	Support	2	X	4	3.5	2	
	Strong support	1	X	2	1.5	1	
		W_I					

Bayesian Methods

Points to be considered

- Attempt to formally summarise totality of evidence (in pre-specified way)
- For some products, no equivalence trial in patients is required => „strong prior“
- Type I error control

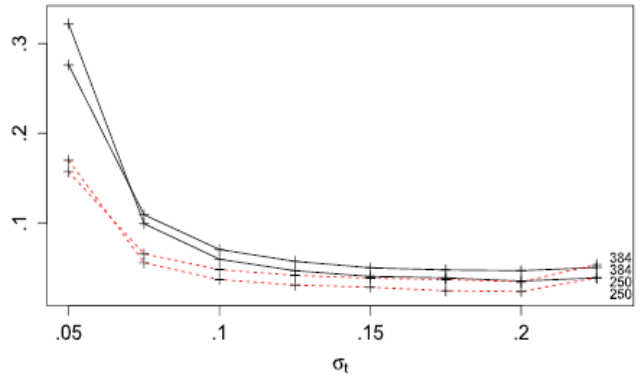

Fig 2, Weiss et al

- Agreement on categorisation of similarity assessment and weighting
- Predictive ability of analytical comparison and phase I study?
- Impact of prior

Biosimilar Development in Rare Diseases

Special methods in small populations?

- Main (feasibility) limitation: Therapeutic eqivalence study

  => need to adhere to same design and endpoints as registrational studies?

- Focus on physicochemical, biological, PK/PD comparability
  - Limited number of batches?
  - Sufficient evidence to conclude that differences observed do not translate into clinically meaningful differences in clinical outcome?
  - Increased rigor? In which sense? Narrower comparability ranges? Simultaneous demonstration of similarity across all critical quality attributes?