Drug Information Association
The Leading Member-Driven, Multi-Disciplinary, Non-Profit Association

2 - 3 November, 19

Brussels Hilt
BRUSSELS, Belgi

NEW ICH GUIDELINE ON CHOICE OF CONTROL GROUP IN CLINICAL TRIALS: THE KEY STATISTICAL ISSUES

PROGRAMME CHAIRPERSON

Karsten SCHMIDT
Chairman EFSPI working group on E10,
Spadille ApS, Denmark

FACULTY

Irwing K. HWANG
Irwing Consulting Group, USA

Bernhard HUITFELDT, EFSPI President,
AstraZeneca R&D, SWEDEN

Meinhard KIESER, EFSPI Council
Willmar Schwarbe Pharmaceutical, Germany

John LEWIS
Medicines Control Agency, UK

Toshihiko MORIKAWA
Takeda Chemical Industries, Japan

Yannick PLETAN, Pfizer, France

Joachim RÖHMEL
Federal Institute for Drugs & Medical Devices, Germany

Barbara VAN ZWIETEN-BOOT
Medicines Evaluation Board, The Netherlands

Mikael ASTROM, EFSPI Council
Active Biotech, Sweden

BACKGROUND

It was decided at the Third International Conference on Harmonisation (ICH3) meeti Yokohama, November 1995, to form an expert working group (ICH E10) on Choice of Cc Group in Clinical Trials. This group should prepare a Step 2 document. This document has released for consultation as a Step 3 document in June, 1999, with a six-month deadlin comments.

The European Federation of Statisticians in the Pharmaceutical Industry (EFSPI) has forn working group to prepare consolidated comments from its ten national groups on the proj new ICH E10 guideline.

The Committee for Proprietary Medicinal Products (CPMP) has also recently prepared a Position on Biostatistical/Methodological Issues Regarding Superiority, Non-Inferiority, and Equivale topic that relates to choice of control group.

The proposed new guideline, draft EFSPI, consolidated comments, and the CPMP pos paper, will be open for discussion in a larger forum at the seminar, which has a faculty cons of ICH E10 expert working group members, EFSPI Council members, and statisticians some of the EU regulatory agencies.

The DIA, in co-operation with EFSPI, organised a similar hot topic seminar on the ICH I Statistical Principle for Clinical Trials based on which, the EFSPI revised its comments t CPMP and successfully contributed to the revision of the E9 final guideline.

ABOUT THE PROGRAMME

The Seminar will commence on **Tuesday November 2, at 10:00** with an introductory se followed in the afternoon by the formation of working groups to discuss various key is Group work will continue the next morning and the working groups will present their r for discussion in plenum and with a panel of regulatory statisticians. The seminar will clo **16:00.** A key feature of the Seminar is its interactive nature.

The submission of written questions to the DIA European Office for processing before Oct **11, 1999,** will help the Faculty, and may receive priority attention at the seminar. Ques received at the Seminar will also receive liberal attention, within the limits of time availab each topic.

WHO SHOULD ATTEND?

Biostatisticians, and others with a keen interest in the design of Clinical Trials, from acad regulatory agencies, the pharmaceutical industry and elsewhere.

A more detailed programme will not be mailed before the Seminar

For further Information, or more details, please contact:
Drug Information Association - European Office
Postfach, 4012 **BASEL** - Switzerland
Tel.: +41 61 386 93 93 - Fax: +41 61 386 93 90 - E-Mail: diaeurope@stepnet.d

NEW ICH GUIDELINE ON CHOICE OF CONTROL GROUP IN CLINICAL TRIALS: *THE KEY STATISTICAL ISSUES*

TUESDAY
November 2, 1999

09:00 Registration & Welcome Coffee

10:00 Session 1
Session Chairperson:
Karsten Schmidt
Spadille ApS, Denmark

Chairperson's Introduction
Karsten Schmidt
Spadille ApS, Denmark

The New ICH E10 Step 3 Draft Guideline on Choice of Control Groups in Clinical Trials
- Background
- Development
- Possible Impact

Yannick Plétan
Pfizer, France

10:55 Session 2
RECENT CPMP ACTIVITIES DIRECTLY RELATED TO THE ICH E10
Session Chairperson:
Mikael Aastrom
Active Biotec, Sweden

Biostatistical/Methodological Issues Arising from Recent CPMP Discussions on Licensing Applications
- Superiority
- Non-Inferiority
- Equivalence

John A. Lewis
Medicines Control Agency, UK

Biostatistical/Methodological Issues Arising from Recent CPMP Discussions on Licensing Applications
- Choice of Delta

Joachim Röhmel
Federal Institute for Drugs and Medical Devices, Germany

12:05 Session 3
MOST IMPORTANT POINTS TO CONSIDER FROM THE PERSPECTIVE OF REGULATORY AGENCIES
Session Chairperson:
Mikael Aastrom
Active Biotec, Sweden

The US-Perspective
Robert O'Neill
Food and Drug Administration, USA

The EU-Perspective
Barbara van Zwieten-Boot
Medicines Evaluation Board, The Netherlands

12:45 Luncheon

14:00 Session 4
MOST IMPORTANT POINTS TO CONSIDER FROM THE PERSPECTIVE OF THE PHARMACEUTICAL INDUSTRY
Session Chairperson:
Mikael Aastrom
Active Biotec, Sweden

The Japanese Perspective
Toshihiko Morikawa
Takeda Chemical Industries, Ltd., Japan

The US Perspective
Irving K. Hwang
Irving Consulting Group, USA

The EU Perspective
Bernhard Huitfeldt
AstraZeneca, Sweden

15:00 FORMATION OF WORKING GROUPS
Session Chairperson:
Karsten Schmidt
Spadille ApS, Denmark

15:20 WORKING GROUP SESSIONS

Coffee placed in the break-out rooms

18:30 RECEPTION

19:30 End of Day 1

WEDNESDAY
November 3, 1999

09:00 Session 5
DISCUSSION OF THE FIVE WORKING GROUP TOPICS, EACH INTRODUCED BY THE WORKING GROUP CHAIRPERSON. THEREAFTER GENERAL DISCUSSION OF THE E10 AND THE CPMP POINTS TO CONSIDER
Session Chairperson:
Bernhard Huitfeldt
AstraZeneca, Sweden

10:15 Coffee Break

10:45 Session 5 cont.

12:00 Luncheon

13:30 Session 5 cont.

15:30 Conclusion to be Drawn from the Discussion
Karsten Schmidt
Spadille ApS, Denmark

16:00 Close of the Workshop

Major comments/needs for clarification from the discussion about ICH/E10 on the DIA meeting in Brussels 99.11.03

By Mikael Åström

Below the points given by the working groups are given. As well comments made during the discussions are given. Some comments may be passed over, some on purpose by me, some because I did not get the point of the comments.

THE SENSITIVITY CONCEPT

- Discussion centered on use of active control to show efficacy.
- Do we understand the problems of active control – Yes, but maybe not the same understanding? KS pointed out a contradictory about the "assay sensitivity" (Section 1.5)
- Do we accept the terminology – No, causing confusions
- Can guideline be clearer – What is necessary to do?
  - Trial sensitivity; External – history of the active control; Internal – design of the study
  - Guidance;
    - External; Seek evidence that the active control shows consistent efficacy in a reliable cohort of well conducted trials (other data sources?, evidence of low, consistent placebo effect))
    - Internal;
      1. New trials are well designed , same patients populations, endpoints -> E9 standards.
      2. Internal support for new trials test treatment are active
- Secondary EP
- Other primary outcome
- Other points;
  - "State of the art" improvements must be permitted.
  - Many situations when external trial cannot be shown but ...
- Make the guideline sharper
- More balance towards non-inferiority trials
- Test for ICH guidelines – does it actually help efficient drug development across the three regions?
- Drug development program
- The role of secondary endpoints to verify assay sensitivity, may help to show that there is quality in the trial

CHOICE OF CONTROL GROUPS AND CORRESPONDING DELTA

- For other working groups
  - The role of placebo
  - Different types of placebo
  - Special types of designs for ...
- Choice of control treatment and patient group
  - Best/optimal control group: criteria? Could the possibility to perform the study in a blind way be a criteria? (Joachim Röhmel; Yes, maybe)
  - Interaction control treatment – patients groups

  - Harmonization achievable? (comments: one comparator, one region (John Lewis); do not do a non-inferiority multiregional trial but a multiregional superiority trial is ok (Robert O'Neill))
  - Different control treatments in one trial?
  - Non-inferiority trials – more ethical?
- Choice of the delta
  - Delta confusion for non-stats.
  - Δ dependent from the size of the effect, no more appropriate if changed (comments: the question is if we can shrink the Δ ?, there is no way to get accept for increasing the Δ (O'Neill).)
  - Δ only (?) be derived from historical / experimental data. This set of trials give evidence for "sensitivity to drug effects" at the same time (comments: if a drug is safe, and shown efficient, i.e. statistically significant, against placebo there is a law in the US saying that the drug should be approved, no matter the effect size. However, if the effect size is to be given in the labeling it should be proven as well (O'Neill))
  - Selection bias
  - Δ should <u>not</u> be dependent on size of related studies or meta-analyses
  - Δ joint effort for clinicians and statisticians
  - harmonization issue, is the same data acceptable in all regions?
- Recommendations: More positive examples for both types of "choices"

STATISTICAL METHODOLOGY

- Incorporate better wording of ptc inte E10 on
  - CI approach
  - Sample size considerations
  - PP and ITT (full analysis)
- Merits of E10 & ptc in terms of marketing applications
  - Only pivotal trial
  - Pooling analyses
  - Clinical development plan (comments: should not be a part of the E10 (Irving Hwang).)
- Superiority
  - From discussion; showing superiority does not mean that "drug is superior". Questions
    - Why not elaborated in ptc?
    - Should "superiority trials" be replaced by "efficacy showing trials"?
- A new "triangular trap", but however the ptc paper will be accepted in all three regions, what happens if not? The EU is going in one direction, do the other regions follow or do they go in their own directions?
- Comment: The E10 should (1) state that it is for confirmatory trials (2) focus more on CI than on hypothesis testing (from the audience).
- 2.5% vs 5%: contradiction between the E9 and the ptc paper. And in the E9 it is stated that for non-inferiority one-sided intervals should be used. However, the probability of covering is not discussed in the E9, which may indicate that the 1-sided interval should be 97.5%. (comment: One sided accepted if stated clearly (Joachim Röhmel).)
- Technical issues
  - No mentioning on the multiplicity issue when switching from superiority to non-inferiority (ptc)

- E.g. if p=0.002 then a 95% CI will exclude zero by a considerable margin. This is not necessarily so, if the size of the trial is huge we may get a low p-value event if the effect size not is large, it is a matter of precision (pct (or E10)).

MODIFIED DESIGNS TO RESOLVE PROBLEMS

- Test/Active control/placebo ("Gold Standard". Comments: not mentioned as gold standard in the E10 (Irving Hwang).)
  - Comment: Should be clear about the objective of the study, is it test vs placebo, or test vs active that is of primary interest (Bernhard Huitfeldt)
  - Add more doses of test: dose response info?
  - Placebo unethical: Skewed randomization
- Factorial design
  - Especially suited for drugs used in combination
  - Could be used in variety of settings with >1 treatment
- Add on study
  - Non-inferiority study cannot be done
  - No information about monotherapy
- Replacement study
  - Standard gradually removed
  - Provides some information about monotherapy
- Early escape (not really a design)
  - Prompt removal of subjects whose condition worsens who fail to respond
  - Criteria for escape to be defined in the protocol
  - Ethical concern about the trial
  - Huge number of patients dropping out (disadvantage)
  - Comment: How do we deal with a patient who are withdrawal – before given their endpoint - from treatment but not willing to participate any more, this may be the case in MS, for example. This should be clarified in the guideline. The message is that we should try to get information even if the information is of no value. (Robert O'Neill)
- Limited placebo period
  - When prolonged placebo exposure not is acceptable
  - May establish assay sensitivity
- Randomized withdrawal (all starts on testdrug, some are switched to placebo in a randomized way)
  - Can be used to study long-term persistence of effectiveness when long-term placebo treatment is not acceptable.
  - Possible extensions to more doses.
  - Only responders will make it to randomization.
  - Carry over?
- General points
  - Acceptability of designs, given limitations?!
  - Feasibility of having placebo in the trial
  - Comment: According to the E10; The priority is to try to do a placebo controlled study, if not possible – try any of the alternative designs, if still not possible, non-inferiority trials may be considered (Bernhard Huitfeldt made the comment, Irving Hwang agreed).

OTHER DISCUSSIONS

- The new "triangular trap", i.e. will the ptc paper be accepted in the US and in Japan?
  - Robert O'Neill: The paper pretty much look like if has been written it. However, there are some process problems.
  - Japan (Dr. Morikawa): The paper is ok.
  - John Lewis: The paper is just the standpoint of the CPMP, will not be released for comments? Detailed papers like this one will be written both in the EU, the US as well as in the Japan.
- Delta references:
  - L. Gould. Some approaches for meta analysis in non inferiority trials, presented at FDA/Industry statistical workshop, Airlington Va., Sept. 30 1999
  - Schmidt, Lau, McIntosh, Cappelleri. An empirical study of the effect of the control rate as a predictor of treatment efficacy in meta-analysis of clinical trials. Stat. In Medicine 17; 1923-1942, 1998
  - Bucher, Guyatt, Griffith, Walter. The treatment results of direct and indirect treatment comparisons in meta-analysis of rando0mized controlled trials. J Clinical Epidemiology, 50, 683-691;1997
  - R. Simon. Bayesian design and analysis of active control clinical trials. Biometrics 55, 484-487, 1999.

//

Conference Report

DIA Workshop
New ICH guideline on choice of control group in clinical trials: The key statistical issues
Brussels, 02-03 Nov. 1999

According to the list of attendees, there were 69 participants, most coming from pharmaceutical industry (50) or CROs (9). There were 6 representatives of authorities, among them O'Neill (US-FDA), Lewis (UK-MCA), Röhmel (DE-BfArM), v. Zwieten-Boot (NL-MEB). Most participants came from France (12), UK and Germany (11), Netherlands (8), Sweden (7), Belgium (5), Denmark (4) and from 6 other countries (11). - The ICH E10 Working Group was represented by the delegates of the pharmaceutical industry from all regions - Plétan (EU), Hwang (US), Morikawa (JA) - as well as by one representative of authorities - v. Zwieten-Boot (EU). The main author of the draft guideline - Temple (US-FDA) - was not present.

The workshop was organised into lectures with sessions
- Contents of the guideline
- Recent CPMP activities related to the guideline (Points to consider, PtC)
- Most important points from the regulatory view
- Most important points from the industry view,

and into working groups
- "Assay sensitivity" and "Sensitivity to drug effects" (Ebbutt, Glaxo Wellcome)
- Choice of control group and corresponding delta (Nowak, ASTA Medica)
- Statistical methodology (Driessen, Organon)
- Design modifications (Jorgensen, Novo Nordisk)

with final plenary presentations and discussions.

Most important background for the E10 draft: Evidently, placebo controlled trials are best suitable to demonstrate efficacy. However, such trials are under criticism (ethical aspects) in many situations now, especially in Japan. Therefore alternatives in terms of active (standard) controlled trials are needed. The guideline clearly states that two validity concepts must be fulfilled in general
- **Assay sensitivity** (AS) or internal validity: The trial must have the property to show differences if they are present. This refers (e.g.) to the choice of the endpoint and to the patient population (in/exclusion criteria).
- **Sensitivity to drug effects** (SDE) or external validity: The selected control treatment must have the property to have reliably shown its efficacy, in general in previous adequate and well controlled trials (i.e. placebo controlled).

AS is not an issue if superiority (either vs. placebo or vs. an accepted control treatment) will be shown; SDE is not an issue if superiority vs. placebo will be shown. In all other cases, especially in all non-inferiority trials, both validity concepts must be seriously considered. In order to demonstrate SDE, it will usually be necessary to perform intensive literature research on the intended control treatment (or on standard drugs belonging to the same drug class as the test drug). - Usually, AS cannot be proven, therefore the trial must fulfil high quality standards.

In addition to the validity concepts, there is an important issue in non-inferiority trials how to define a margin ("delta") that guarantees that the test drug will still be statistically superior to placebo. This margin must be scientifically explained and is best derived from the set of trials selected for SDE by performing a methodologically sound meta-analysis. It is often proposed to define a delta of 50% (or smaller) of the standard drug's effect over placebo. In any case, the delta must be negotiated with the authorities. There may be situations where the determination of the margin must be "traded off" with other benefits (e.g. safety, pharmacoeconomic); this seems unsolvable using statistical methods.

There are many other design variations described in the guideline. But this does not give any way out: The reduced size of placebo patients does not solve any ethical problem and if a placebo is incorporated in whatever way, the validity problems are of minor interest. Interestingly, it was clearly stated that the 3-arm trials (test, standard, placebo), are by no means a "golden" standard: They incorporate a lot of problems because of the many possible tests. What use has a superiority "test vs. placebo" if the superiority "standard vs. placebo" fails (missing AS)? A non-inferiority "test vs. standard" does not help if there is not a simultaneous superiority "test vs. placebo". Additionally, the multiplicity problem must still be discussed. - Even substituting placebo by dose steps of the test drug will lead to the ethical problem for the smallest "non-effective" dose.

A special discussion point was the exclusion of drop-out from efficacy analysis, especially if they give information on efficacy (e.g. placebo drop-out). While drop-out should not be included if there is no endpoint information available (O'Neill), there is a debate on the appropriate endpoint definition. Time-to-event is definitely a good solution. (Remark: There was no discussion on "response" definitions which may be a combination of a dichotomised measure and the drop-out phenomenon itself.) - This issue is important for allowing (reinforcement) "early escape" by protocol. A follow-up (at the planned end) of such cases was suggested (O'Neill). Phrases like "drop-out from study" must be avoided.

It was very interesting that everybody, including the representatives from the authorities, had a "bad feeling" against the draft. This starts from editorial aspects (hard to understand, triple negations, new and uncommon definitions as the validity terms) and ends with the blame that the draft does not help in the harmonisation process. Although almost everything is correct (like a textbook), there are too many aspects where the draft refers to "regional specifications".

Important was the presentation (Lewis) of a brand-new CPMP PtC on biostatistical/methodological issues "Superiority, non-inferiority and equivalence" which is to be commented until Dec. 1999. Remarkable is the discussion of a post-hoc change of the hypotheses ("switching the objectives"). While the interpretation of a non-inferiority trial as a superiority trial gives no problems, there are still open issues for the other direction. Even if there is an a-priori (prospective) definition of the "delta" margin in the protocol, there are unsolved problems of sample size planning and multiplicity in this case. A retrospective widening of confidence intervals (as margins) is not acceptable in any case. Between lines (?), the different role of itt (full analysis set) and pp is described: In non-inferiority trials pp and itt analyses are (at least) equally important (put into question by O'Neill), while itt is the more important analysis in superiority trials. This is of special relevance for the case of switched objectives.

Finally, there were only few concrete ideas for the improvement of the E10 draft, especially in removing the "negative" look on non-inferiority trials. EFSPI (European Federation of Statisticians in Pharmaceutical Industry) will prepare a joint position paper to the draft. Probably on 22 Nov. 1999, an EFPIA (European Federation of Pharmaceutical Industry Association) meeting of their Expert Working Group will be in Brussels. (Remark: I am invited to participate as delegate of the VFA.)

Some interesting quotations:
- "Guidelines are not only for harmonisation, but also for education." (v. Zwieten-Boot)
- "E9 tells what should be done, E10 how it should be done." (v. Zwieten-Boot)
- "E10 is too early for efficacy **and** safety evaluation (benefit/risk)." (O'Neill)
- "The guideline is not easy to understand." (O'Neill)
- "We do understand the guideline. But do we understand the same?" (Schmidt)
- "Authorities look at both, estimates and confidence intervals (test results)." (Lewis)
- "Clinical significance is often non-assessable: A difference is a difference." (O'Neill)
- "Dose response comparisons for test and standard drug are interesting." (Röhmel)

Horst Nowak

- Summary of my working group (Choice of control group and corresponding delta) included.
- Workshop material availble at me.