INTERNATIONAL SOCIETY FOR CLINICAL BIOSTATISTICS

Chairman, Sub-Committee on Statistics in Regulatory Affairs:
Dr. Jørgen Seldrup, Clinical Trials and Epidemiology Research Unit, Ministry of Health, 10 College Road,
Singapore 169851, Republic of Singapore
Telephone: +65 220 1463, Fax: +65 220 1475, e-mail: jorgen@cteru.gov.sg

---

Comments regarding National Institute for Clinical Excellence: "Revised Guidelines for Manufacturers and Sponsors of Technologies Making Submissions to the Institute", First Draft, 12 October 2000.

Major Points

4.1.3. Data Sources

The recognition that clinical trial data will usually have to be supplemented by data from other sources and that modelling approaches will be necessary is realistic and welcome (Rittenhouse, 1997). As mentioned in the last paragraph caution to avoid biases is important. This is particularly true when soft end-points such as quality of life are used in open trials.

4.4.4. Forms of Analysis (last paragraph)
It has been well understood in statistical circles for a long time that power is relevant to the planning of clinical trials, but not for interpreting them (Cox, 1958, pg 161). The power approach, for example, has been abandoned in the context of bioequivalence, where instead it is necessary to show via confidence intervals that treatments are equivalent. The issue of non-inferiority is addressed in various ICH (International Conference on Harmonisation) guidelines in particular ICH/393/96 topic E9, Statistical Principles for Clinical Trials (section 3.3.2) and ICH/364/96 topic E10, Choice of Control Group in Clinical Trials. This is also addressed in a recent CPMP [Committee for Proprietary Medicinal Products document CPMP/EWP/482/99 'Points to Consider' on Switching between Superiority and Non-inferiority (Adopted July 2000).

4.6 Comparisons.

The statement that the main comparator should be the most frequently used treatment in the patient group in question is too restrictive. In some populations the treatment will be an innovation. For example, ethical considerations may require that all patients get the standard therapy and that the new therapy is compared as an "add-on" to placebo.

Also, it would seem logical that as long as a number of current rival treatments are reimbursed, an economic evaluation could be justified by choosing any of them as a comparator, not just the "gold standard".

This guideline should give a reference to ICH/364/96 topic E10, Choice of Control Group in Clinical Trials

4.8 Generalisability.

All clinical research involves a degree of artificial abstraction. The general assumptions of clinical trials are that internal validity is the most serious concern, but that if (approximately) additive measures are chosen

the results will apply (approximately) on that scale to target populations. The invitation to explain the implications of extrapolation is welcome, but it will be unfortunate if an unrealistic fetish of representativeness is made. Carefully chosen clinical measures are more likely to transfer from trial to target population than are costs. This reinforces the necessity of modelling approaches.

4.9.1. Reporting of results

The CONSORT statement is (of necessity) vastly inferior in scope and detail to the various ICH guidelines (for example ICH E9 (International Conference on Harmonisation, 1999)). These guidelines should not be referred to only for planning.

4.9.3. Risk estimates

The choice of outcome measures is disappointing. Number needed to treat, NNTs are not wise choices for summarising results because they do not generalise well from one population to another (Hutton, 2000; Senn, 1998; Smeeth, et al., 1999). Even relative risks are inferior to odds ratios under many circumstances.

There is a danger of confusing the issue of reporting and summarising trials with that of applying the results in practice. For example for a given level of background risk an odds ratio can be converted into an NNT, but whereas the odds ratio may well be roughly generalisable from one population to another, the NNT is unlikely to be. This suggests a strategy of summarising trials (for example in a meta-analysis) using odds ratios and then converting these results via appropriate modelling strategies to target populations.

4.9.4. Subgroups

The paragraphs here are rather contradictory. It is not at all clear what the guideline is saying.

The first paragraph seems to imply that a) sub-group analysis should be carried out if it is believed that cost-effectiveness may vary, particularly for high risk patients b) that tests for interaction rather than within stratum treatment comparisons should be performed and c) adjustment for multiplicity should be made. The optimal strategy, therefore, for a sponsor fearing lack of efficacy in certain target subgroups is to create as many cross-classifications as possible, test several times for interaction and adjust using the Bonferroni correction.

It should be noted that in the case of zanamavir (Relenza), NICE did the opposite of b). "Due to the limited numbers of "high risk" patients (particularly the elderly and those with cardiovascular disease, asthma, chronic obstructive pulmonary disease, or immunosuppression) that have been treated with zanamivir (Relenza) in clinical trials, the Institute has not found it possible to conclude that the product reduces the frequency of serious secondary complications in these groups of patients." (From NICE judgement on zanamavir). However, there did not seem to be any difference between the high risk group and the rest regarding efficacy, so the Institute's judgement appears to be based on a lack of *provable* efficacy within the particular group. In other words, based, essentially on the inspection of a sub-group P-value.

It seems here that the authors of the guidelines are being inconsistent. On the one hand, they wish to avoid sponsors claiming efficacy on the basis of significant subgroup results. On the other hand they desire them to show significant subgroup results. The latter requirement, however, is well-known to experienced trialists to be unrealistic. In many indications, it is not reasonable to expect sponsors to be able to power trials to show efficacy in all sub-groups. The basis of drug development in general is that results, provided they are based on well-controlled studies can, with the aid of reasonable judgement, be projected onto other populations.

<u>Minor Point</u>

The word "data" is treated as a plural in some places (for example 4.1.1) and as singular in others (for example 4.7 foot of page 18). Consistency helps to promote readability. Plural is preferable.

References

COX, D. R. (1958)
Planning of Experiments, New York: John Wiley.

HUTTON, J. L. (2000)
Numbers needed to treat: properties and problems (with comments)
Journal of the Royal Statistical Society A, **163**, 403-419

INTERNATIONAL CONFERENCE ON HARMONISATION (1999). Statistical principles for clinical trials (ICH E9)
Statistics in Medicine, **18**, 1905-1942

RITTENHOUSE, B. E. (1997)
Exorcising protocol-induced spirits: Making the clinical trial relevant for economics
Medical Decision Making Canada., **17**, 331-339

SENN, S. J. (1998)
Odds ratios revisited
Evidence-Based Medicine, **3**, 71

SMEETH, L., HAINES, A. & EBRAHIM, S. (1999)
Numbers needed to treat derived from meta-analyses--sometimes informative, usually misleading [see comments]
BMJ, **318**, 1548-51