Clin. Eval. 26 : 475-485, 1999. Jan.

The Third Party Guarantee System in 798 Double Blind Randomized Controlled Trials in Japan

ABSTRACT

Naokata SHIMIZU[1)2)], Yorio SATOH[2)], Masanao KURIHARA[2)], Kiichiro TSUTANI[3)]
1):Teikyo University 2):Controller Committee 3):Tokyo Medical and Dental University

The third party guarantee system of clinical trials, which is called as "The Controller's System", is a unique and original system developed in Japan. The Controller guarantees qualities of every aspect of the clinical trial, including design, blindness, raw data, statistical analysis, interpretation and publication. In the period of 1971 to 1998, we participated in 798* double blind randomized controlled trials (DB-RCT) in 12* drug categories, including 241* placebo-controlled trials. In this presentation, we present this unique system and the results.

The Controller, independently from investigator and sponsor of the trial, participates in and guarantees the quality of every aspect of the trial. Detailed operating procedures to maintain the reliability such as unique "double controller system" will be presented.

The names of drugs chosen as control, the numbers of subjects who were included or excluded for analysis, methods of statistical analysis, results of those trials and their meta-analysis data in different drug categories will be presented. For example, the results of meta-analysis and global improving rate of controversial drug categories such as drugs for cerebrovascular diseases, consisted of 22 trials against placebo revealing that superior results to placebo in 10 trials, inferior in 2 trials and no difference in 10 trials.

Since the new Good Clinical Practice(GCP), on the basis of ICH-GCP, was introduced also in Japan in 1997, almost all responsibilities of quality assurance are left to the sponsor of the trial. The roll played by the third party at this time will be reevaluated in the presentation.

(*: Figures are slightly changed according to the newest data.)

The Controller Committee consists of those members of internist, psychiatrist, gynecologist, ophthalmologist, pharmacologist, clinical pharmacologist, lawyer, statistician and pharmacist. The Controller Committee publishes the quarterly Journal *Rinsho Hyoka* (*臨床評価*、*Clinical Evaluation*) which content is available though our home page "*http.//www.sphere.ad.jp/cont*".

Introduction

The contents of this print material distributed today are mainly excerpts from a publication by authors who are also the members of the presentation group (Hiroe Tsubaki, Toshiharu Fujita, Yorio Satoh: On the Patient Oriented Clinical Evaluation In Japan and Its Statistical Aspects, appeared in Proceedings of the Institute of Statistical Mathematics, vol. 46, pp.97-115, 1998). The purpose of this document is to foster understanding about the objectives the group would like to present. The following are the five main points which we wish to emphasize in the presentation, including the discussion of multiplicity and equivalency.

1. Principles of hypothesis validation (eliminate conclusions through exploratory analysis which could utilize multiplicity.)
2. Rigorous application of equivalency inference
3. Emphasize the significance of pragmatic trial
4. Encourage to analyze and not eliminate incomplete cases whenever possible
5. Establish and perform validation inference on three main analysis parameters associated with usefulness, efficacy and safety.

Changes in clinical evaluation at ICH and Japan

The activities under ICH take into consideration the producers' "feasibility" to provide "excellent pharmaceutical products at a faster speed for patients around the world," and efforts have been made to "implement scientific clinical trials", which would also be the most economical strategy for producers. ICH rules value tests that prove superiority in clinical trials in order to demonstrate convincing data to establish superiority, such as "superiority over placebo", "superiority over active control drug", and "to demonstrate dose-response relationship". They are directed towards careful selection and refinement of endpoints in advance, or towards reduction of the number of participating institutions to increase the number of cases per institution. They have a distinct characteristics of an "explanatory trial" or a "scientific trial" with an objective to perform "scientific" validation of hypothesis under controlled conditions.

In the past, our group has valued the "usefulness evaluation" and "technical evaluation", to assess behavior of drugs that are newly introduced into the society, from the standpoint of patients in the clinical evaluation practice in Japan. Are these practices no longer necessary? According to the ICH statistics guideline, these are the "discussion of generalizability", but an alternative way to implement these evaluation as

post marketing survey after approval is also presented. In addition, although the ICH statistics guideline demand two or more validation tests, tests with different characteristics rather than repetition of similar tests should be acceptable. Pragmatic tests for the purpose of technical evaluation is strongly desirable as a Phase Three trial.

Usefulness evaluation and general evaluation from the standpoint of patients

Many argue that the poor quality of Japanese clinical trials derive from the usefulness evaluation and the small number of cases per institution and that improvements are necessary. But is it really so ?

The statistical analysis guideline stresses the necessity of multiplicity adjustment and validation of hypotheses. On the other hand, it developed a side effect that is to neglect analysis of ancillary items. This would augment a tendency in which clinicians cannot obtain from clinical trials diverse information necessary to adequately evaluate new drugs, and such situation is likely to lead to a reversal phenomenon and a stagnant thinking, in which the responsibility to judge the usefulness of a new drug based on the data from clinical trials lie on the part of statisticians. This would undermine the ultimate goal of clinical evaluation.

Merely relying on the philosophy of the three principles in the Fisher style scientific test plan (repetition, randomization, local control) as the bases is not efficient in order to discuss down-stream reproducibility. In general, a down-stream reproducibility of a product function means that the product is free of any functional troubles even when put to use under diverse conditions that are normally not tried in experimental stages.

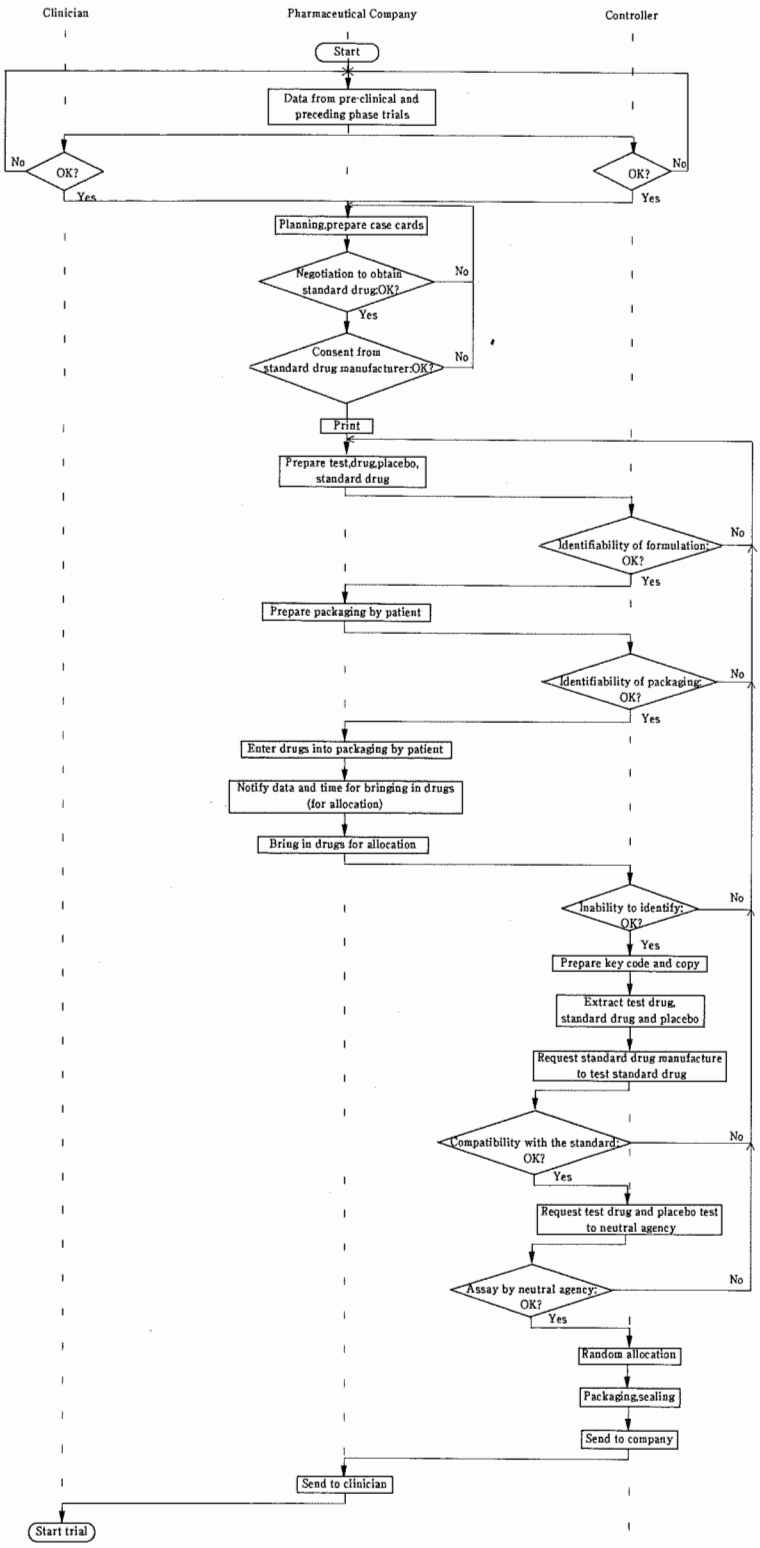
Even if one yields and take the side of Fisher's classical test plan, still the practical objective in the principle of a local control is that it would be necessary to actively introduce a large number of institutions as block factors to reduce the weight the complete randomized trials take within experiments, in order to ensure the universality of the conclusion. Collecting information from multiple institutions merely means a cluster extraction (normally it is a significant extraction) of patient population within an institution, and unless the inter-cluster variation is sufficiently smaller compared to the intra-cluster variation, it is preferable to have larger cluster numbers within the same sample size from the viewpoint of survey and measurement accuracy. This is the fundamental recognition of a survey design (Cochran 1961).

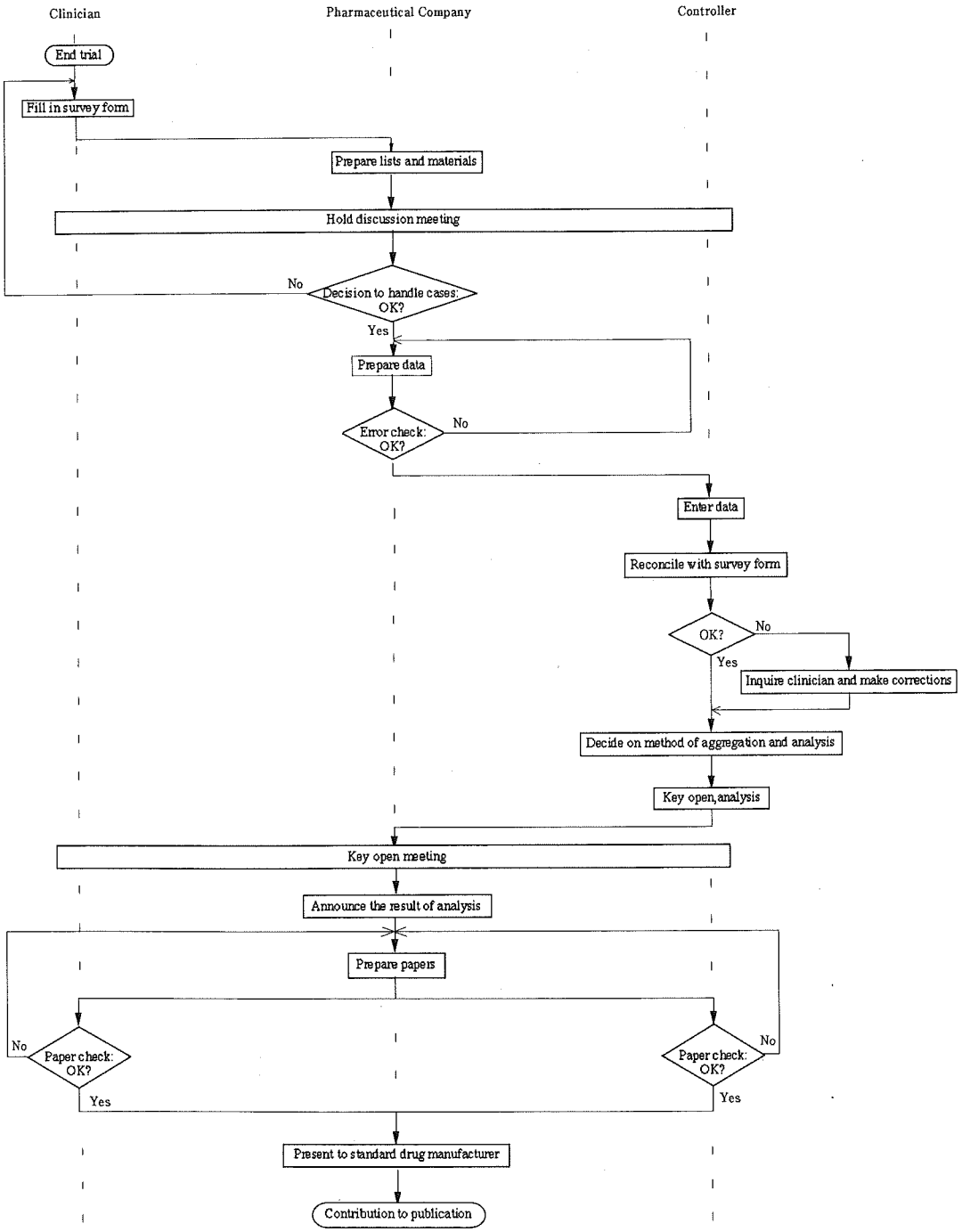
In order to perform quality engineering experiments, in other words to mimic actual conditions, it is more efficient to perform tests based on a test plan which intentionally introduce sufficient noise factors. From the viewpoint of measuring clinical usefulness of a therapeutic method, a scheme which introduce more blocks are definitely superior over trials performed at smaller number of institutions adopting precision science methods. Because, a "statistical method for scientists" is merely to perform efficacy assay by representing institutions as denominating block effects, and this constitutes an efficacy assay based on "repeatability variance". Whereas in a "statistical method for technicians", noises that exist in clinical settings of trials at multiple institutions are taken in as noise factors (indicated factors), and efficacy is evaluated on the reproducibility variance, based on cross reaction between these noise factors and efficacy (control factors). The "statistical method for technicians" based on reproducibility variance leads to direct evaluation of efficacy and usefulness in the market than the "statistical method for scientists" which is based on the repeatability variance. Naturally, not all noise factors are necessary in clinical trials. There are many noises such as lack of evaluation reliability that need to be reduced through appropriate standardization. However, at least during the period of expected use for a new drug, a philosophy to ignore noises that would exist in the clinical environment or eliminate them through statistical analysis are the acts that the users (patients) would never accommodate.

Quality Assurance

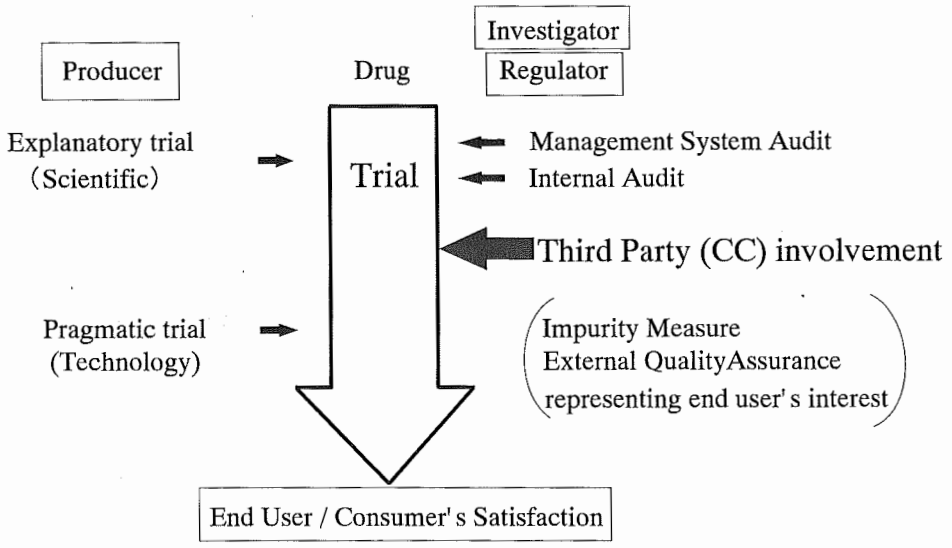
In a sense it is an interesting experiment that the new Japanese GCP based on the ICHGCP obligated the "producers only" to bear the sole responsibility of the entire course from design and implementation of clinical trials, structuring of quality assurance system, to the external quality assurance. Along with this, in Japan, the official publication of reports from clinical trials is no longer obligated, and there has been a shift in roles from the third party controller to the advisors to manufacturers. There is a possibility that the Japanese quality assurance system which in principle surpass that of the Western system be dismantled without an empirical comparison.

Would it be possible to guarantee impartiality in the Japanese clinical evaluation through the quality assurance by producers (pharmaceutical companies) such as monitoring and auditing regulated under the new GCP, and through inspection by regulatory authorities? It is certain that in many cases these systems would be useful in enhancing qualities of clinical trials


Clinician
Pharmaceutical Company
Controller
Start
Data from pre-clinical and preceding phase trials
No
OK?
Yes
OK?
No
Yes
Planning,prepare case cards
Negotiation to obtain standard drug:OK?
No
Yes
Consent from standard drug manufacturer:OK?
No
Print
Prepare test,drug,placebo, standard drug
Identifiability of formulation: OK?
No
Yes
Prepare packaging by patient
Identifiability of packaging: OK?
No
Yes
Enter drugs into packaging by patient
Notify data and time for bringing in drugs (for allocation)
Bring in drugs for allocation
Inability to identify: OK?
No
Yes
Prepare key code and copy
Extract test drug, standard drug and placebo
Request standard drug manufacture to test standard drug
Compatibility with the standard: OK?
No
Yes
Request test drug and placebo test to neutral agency
Assay by neutral agency: OK?
No
Yes
Random allocation
Packaging,sealing
Send to company
Send to clinician
Start trial


Clinician
Pharmaceutical Company
Controller
End trial
Fill in survey form
Prepare lists and materials
Hold discussion meeting
No
Decision to handle cases: OK?
Yes
Prepare data
Error check: OK?
No
Enter data
Reconcile with survey form
OK?
No
Yes
Inquire clinician and make corrections
Decide on method of aggregation and analysis
Key open, analysis
Key open meeting
Announce the result of analysis
Prepare papers
No
Paper check: OK?
Yes
No
Paper check: OK?
Yes
Present to standard drug manufacturer
Contribution to publication

Conceptual Diagram of Drug Development and Evaluation from the Standpoint of Quality Assurance for End User



OBJECTIVES

1. **To delineate the involvement and experiences of a third party (e.g. Controller Committee/CC) participation in the 798 DB-RCTs conducted in Japan;**
2. **To expound the importance of a third-party guarantee system (using the CC model) to ensure the quality and fairness of RCTs.**

BACKGROUND

- In the early 1970s, several Japanese clinicians pooled together to discuss and sincerely address issues relating to clinical evaluations in Japan. This led to the establishment of the Controller Committee (CC) in 1972.

- Three fundamental principles of clinical evaluation have been established by the CC, i.e.:
  - Least biased information principle;
  - Patient-oriented evaluation principle; and
  - Standardized clinical trial and evaluation procedure principle.

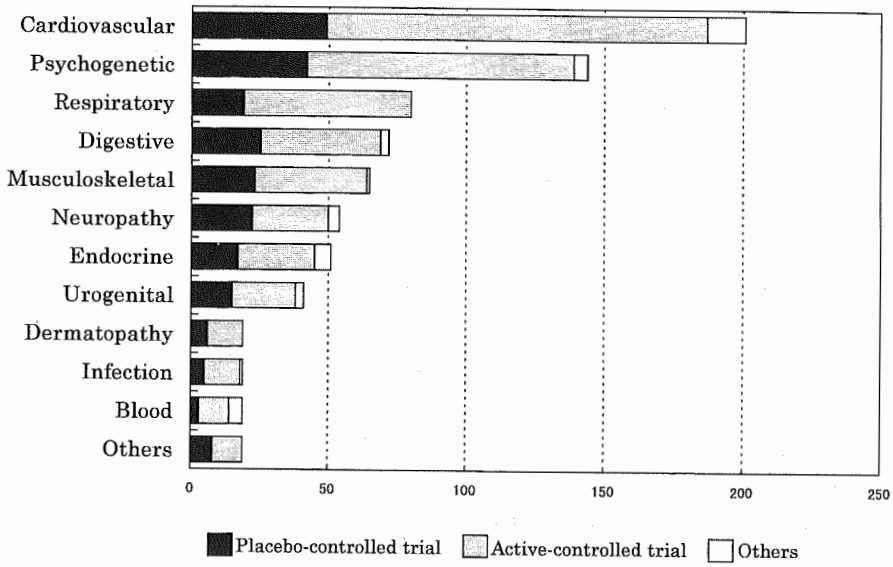
- 798 DB-RCTs were conducted in Japan following these three principles.

- This unique and original Japanese methodology in clinical evaluation is introduced here as well as a discussion of our views on the ICH-GCP.

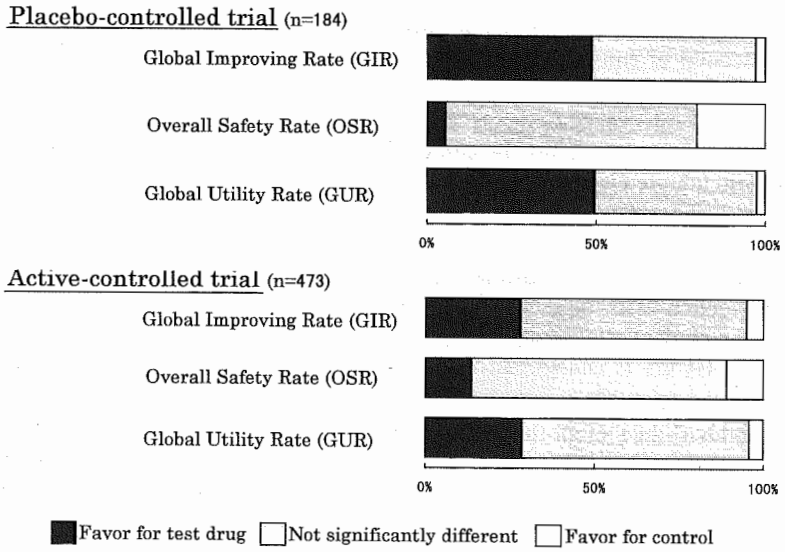
THE CONTROLLER COMMITTEE (CC) SYSTEM

- A conformity assessment system to eliminate any possible bias which may be induced by the intentional intervention of an interested party (e.g. sponsor) during the different stages of the clinical trial.
- A quality assurance system conducted by an NPO, comprising of expert clinicians independent from both the sponsor and trial investigator, using the "double controller system", external auditing of protocol, etc.
- An external auditing system which is different from the internal auditing system contained in the 1996 ICH-GCP and the new J-GCP (1997).

Area of Randomized Clinical Trials involving the Controller's Committee (total n=798)



Test Drug vs Control (Three Global Assessment) (total n=657)



THE CC SYSTEM PROCEDURES

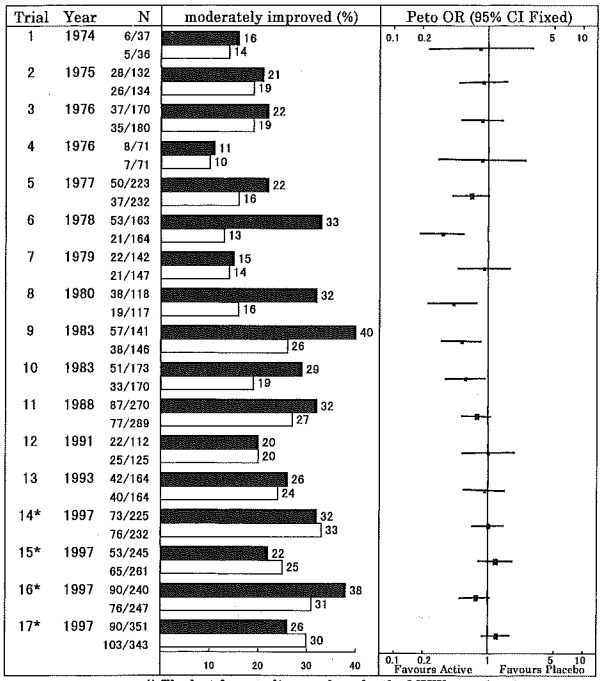
The Controller Committee uses strict procedures to ensure the quality of and fairness in the conduct of clinical trials:

- **Review of the trial protocol from the viewpoint of consumers/patients;**
- **Guarantee in distinguishability between control drug and test drug;**
- **Random allocation and ensuring blindness of the trial until key break;**
- **Sampling and testing the test drug to ensure that it meets formulary standards;**
- **Review the validity of the statistical plan;**
- **Review of both included and excluded cases for final statistical analysis under blinded conditions;**
- **Collect all case report forms (CRFs) including results of laboratory tests for data analysis and maintain data archive for future analysis/evaluation;**
- **Conduct statistical analysis on designated endpoints, ie, GIR, OSR, GUR;**
- **Review draft trial paper/report before publication to attest to the contents of the report.**
- **Publication of results of all trials conducted including those trials with negative results to avoid bias in publication.**

CONCLUSIONS

1. **DB-RCT should be conducted keeping foremost in mind the interest and protection of end-users (consumers/patients) and not the protection of the sponsors interest.**
2. **To avoid intervention of interested parties to the investigation, external quality assurance should be conducted in addition to the internal auditing system adapted in the ICH-GCP.**
3. **In the Phase III study, the Global Usefulness Evaluation (GUE) of each case measured by the individual investigator-in-charge as well as the patient is important and of primary value in determining customer satisfaction after product marketing. For this purpose, the broad multi-center clinical trial is more meaningful in measuring impurities since the Phase III CT is considered as a pragmatic trial or technology assessment and not a scientific experiment or explanatory trial.**
4. **Individual data of all the 798 trials conducted were generated under same standards and archived in EDP form to allow for secondary and meta-analysis or systematic review such as those used in the Cochrane Collaboration.**

Efficacy of Drug Treatment for Chronic Cerebrovascular Disorders in Comparison with Placebo



*) The last four studies are done for the MHW reevaluation program

Test Drug ☐ Placebo

Implementation of publication policy including "negative" trial (total n=642)