

NICEATM informed the panel of the current international validation efforts of the LUMI-cell transcriptional activation assay for agonists and antagonists, beginning now, under the auspices of ICCVAM.

6.5 Action Item: CERi to conduct an internal audit of data transcribed.

This was done according to GLP; one error was identified in Table 13, which has now been corrected. A modified Table 13 will be attached in the final report.

7. Action Item: Proposed methods for estimation of between- and within-run (laboratory) variation.

Agreement on future plans on the revision of and addition to the analysis. Dr Yutaka Aoki (US EPA) gave a presentation with a focus particularly on a weighted average approach for assessing between- and within-run (laboratory) variation and the calculation of standard deviation (SD), with a view to refine the estimates of the various sources of variability that contribute to differences in response. Two macros were also included for the panel participants to experiment with. This presentation can be included as an appendix in the final summary, together with the response of Joe Haseman (see item 10.2 and appendix 2 attached). Further discussion on this will be useful. Dr Sebastian Hoffman, the ECVAM statistical expert has been sent Dr Aoki's presentation for his informal input. It was understood that with other ongoing validation work within ECVAM, SD's are favoured over CV's.

New Action item: NICEATM to contact consultant statistician Dr Joe Haseman, and The Secretariat to contact Dr Sebastian Hoffman providing Dr Aoki's presentation for consideration. **Done.**

Dr Aoki offered to conduct statistical estimations as described, on the CERi data probably in mid June 2006 (but this is not finalised), for the peer review.

New Action item: Dr Aoki to conduct statistical estimations of between- and within-run (laboratory) variation provisionally in June 2006.

8. Cytotoxicity queries: Provision of the criteria for when cytotoxicity is evaluated and how the data are interpreted, together with the provision of such data with respect to the reproducibility of the cytotoxicity assay when conducted at a different time and using a different (but related) cell line. (ICCVAM request following the 2nd teleconference.)

CERi explained that generally, when the cell viability is below 80% of the solvent control, the test concentration is regarded as a cytotoxic concentration. When the tested concentration is regarded as a cytotoxic concentration, the data at that concentration is excluded from the antagonist data analysis. CERi do not have data on the reproducibility of the cytotoxicity assay at this point.

9. Next steps:

9.1 Timelines for pre-peer review report.

By the end of May 2006, the draft report will be circulated to the panel participants, with finalisation of report by mid-June 2006. Further related activities are discussed in agenda item 5.

9.2 Date and time for statistics teleconference?

It was agreed that this is not necessary at this time. However following the comments received from Joe Haseman (see item 10.2), statistics consultant to ICCVAM, perhaps a statistics teleconference and/or further email correspondence would be of value?

10. Any other business

Participants expressed their gratitude to CERi and Dr Aoki for their extensive hard work in providing all the requested data and agenda items.

CERi and the OECD Secretariat expressed their gratitude to the participants for their expertise and constructive input.

11. Adjournment

The call leader adjourned the meeting at 3.30pm (CET).

10. Subsequent developments

Following the teleconference, additional observations with respect to edge effects were discussed by Drs Jun Kanno and Miriam Jacobs. There can be a number of plate effects one might usefully consider in addition to those mentioned during the teleconference, e.g.

1. There can be effects due to cell respiration and metabolism that can be affected by cell number in each well, such that the greater the cell density required by a protocol, the more unhealthy or depleted the cells in central wells might be, compared to those at the edge.
2. Optical differences in position of the different wells of the plates can affect the luminosity readings by a plate reader (as well as observation by the naked eye).
3. Stacking of plates: effects on cell metabolism have been observed in plates at the bottom of the pile of stacked plates when a large number, i.e. more than 8 plates have been stacked on top of one another in the incubator.

On 25 May 2006 the Secretariat received:

10.1 NICEATM draft comments on the CERI submission (appendix 1) and

10.2 Joe Haseman's (Consultant statistician to ICCVAM) comments on the statistical analyses proposed by Dr. Aoki (appendix 2).

Appendix 1

Draft NICEATM Comments on Studies Conducted by CERI to Support the Validation of the hER-HeLa-9903 Estrogen Receptor (ER) Transcriptional Activation (TA) Test Method

Our comments are based on information CERI has provided in their report entitled, “Draft Pre-Validation and Inter-Laboratory Validation Report of the Human Estrogen Receptor Mediated Reporter Gene Assay”, and other supporting materials, including those used to present information that CERI has provided at the request of the OECD Preliminary Validation Assessment Panel. Our assessment of the provided information is based on relevant information provided in Section VII of OECD Guidance Document No. 34, which recommends and defines the components of a new test method submission. Our assessment of the hER-HeLa-9903 ER TA test method protocol is based on the minimum procedural standards (we now call these essential test method components) recommended by ICCVAM¹ and based on the deliberations of an ICCVAM international expert panel on ER and androgen receptor binding and TA assays that met in May of 2002. Our evaluation of the substances used to evaluate the accuracy and reliability of the hER-HeLa-9903 ER TA test method is based on the ICCVAM list of recommended reference substances for ER binding or TA test methods².

Our comments are organized under the major headings in Section VII of OECD Guidance Document No. 34 as follows:

Introduction and Rationale for the Proposed Test Method

Reports and supporting materials address the rationale for the CERI ER TA test method, as specified in this section of the Guidance Document, but discussions regarding the specific limitations of the test method could be usefully expanded.

Test Method Protocol Components

A test method protocol has been provided, as specified in this section of the Guidance Document, but this is the protocol that was used for the experiments that involved multiple laboratories only. It is stated in the text that the in-house protocol was similar but the protocol followed throughout and any modifications and the rationale for those modifications needs to be included. For example, in the interlaboratory study, estradiol was tested over multiple concentrations but in the in-house studies, it was tested at only a single concentration. The rationale for this difference should be provided.

In addition, in terms of the test method protocol, the highest concentration of substance tested was 10 μ M, not the 1 mM recommended by the ICCVAM international expert panel and ICCVAM (see footnote 1). We appreciate that not all substances can be tested up to this concentration (due to solubility or excessive

¹ “ICCVAM Evaluation of *In Vitro* Test Methods For Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays” (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

² “ICCVAM Evaluation of *In Vitro* Test Methods For Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays” and the 2006 Addendum to this report (available at <http://iccvam.niehs.nih.gov/methods/endocrine.htm>).

cytotoxicity) but the purpose for using this limit dose is to detect even very weak ER agonists or antagonists. Thus, at least some of the substances classified as negative by CERI have not been adequately tested (this was demonstrated in the data set provided by CERI for the last conference call) while others may have been adequately tested if solubility or cytotoxicity data can be provided to support the highest concentration tested.

There seems to be a lack of information in regard to the rationale/justification, criteria for use, and reliability for the cytotoxicity evaluation, which were conducted using the same basal cell line but with a different plasmid construct as a separate experiment. From verbal discussions, it appears that CERI does not feel a cytotoxicity evaluation is needed for the agonist tests. This issue needs to be formally discussed in their submission.

For use as a screening assay for ER or AR activity, it is critical that a TA test method evaluate for antagonist as well as agonist activity. Except for the intralaboratory repeat testing of three substances, an evaluation of the ability of the CERI ER TA test method to identify ER antagonists has not been provided. Furthermore, the antagonist protocol used in the testing of these three substances had no concurrent positive control, and did not use a reference standard with a full dose response curve as is done in the CERI agonist protocol. We appreciate the desire to move ahead with the agonist version of the test method independent of the antagonist version but wish to point out that a negative ER agonist study is virtually worthless without knowing whether or not the test substance binds to the ER and/or demonstrates antagonist activity. We do not agree with CERI's premise, stated in the most recent OECD teleconference, that the antagonist protocol is similar enough to the agonist protocol to be considered as validated in the same manner. We urge that the current ER antagonist protocol be modified to include appropriate positive controls and that further validation studies using this protocol be completed before peer review.

The protocol needs to include a discussion about potential "edging effects", and how to identify if the outside wells on the 96-well plate can be used because such effects are not detected under the experimental conditions used by a specific laboratory.

Characterisation and Selection of Substances Used for Validation of the Proposed Test Method

To facilitate validation of ER TA assay, ICCVAM compiled a list of 78 recommended reference substances. ICCVAM recommends that these substances be tested in a phased manner, with a minimum of 53 substances being tested across at least three laboratories. The remaining 25 substances are recommended for testing once in one laboratory or divided among two or more laboratories.

Our evaluation of the data submitted indicates that CERI tested a total of 56 substances, although only 10 were tested across multiple laboratories. Seven of these 10 substances are on the ICCVAM list and the remaining three have similar ER activities to other ICCVAM substances recommended for interlaboratory testing and could be considered as replacements for these.

Therefore, to meet ICCVAM recommendations, 43 additional substances from the ICCVAM recommended list or their equivalents would require further interlaboratory testing.

CERI tested 12 of the remaining 25 substances on the ICCVAM list that do not require interlaboratory testing at least once, leaving an additional 13 substances from the list or their equivalents that would require further testing.

Also, substances are not classified according to product class and only the 10 substances tested across multiple laboratories are classified by chemical class. These 10 substances represent 6 chemical classes

compared to the 15 chemical classes represented by those substances recommended for interlaboratory testing by ICVAMM (a total of 22 chemical classes are represented by the ICCVAM recommended list of 78).

In Vivo Reference Data Used to Assess the Accuracy of the Proposed Test Method

The comparison of experimentally derived results from ER TA agonist and immature rat uterotrophic studies conducted at CERI using 50 substances adequately supports the accuracy of the proposed ER TA agonist test method.

Testing all 78 reference substances would not only allow for a better characterization of the reliability and comparative sensitivity of the CERI test method versus other Tier 1 assays but also increase the likelihood that *in vitro* tests might be developed that could be used to reduce animal use in endocrine disruptor (ED) testing.

Test Method Data and Results

Results and data from prevalidation and interlaboratory studies conducted by CERI to support the validation of their hER-HeLa-9903 ER TA agonist assay have been provided, but much of this was not provided in the CERI draft validation report but rather at the request from the OECD preliminary validation assessment panel. It is assumed that the requested results and data will be included as appropriate in the appendices of the final validation report from CERI.

Test Method Relevance (Accuracy)

Because this test method is to be used as a Tier 1 screening assay (at least in the United States), there is no need for an evaluation of the ability of the test method to predict *in vivo* endocrine disruptor effects. However, such data are welcome and would allow better characterization of the ability of *in vitro* test methods such as this to reduce animal use in ED testing. The comparison of CERI derived ER TA results with ICCVAM published ER TA results for 46 substances is appropriate.

Test Method Reliability (Repeatability/Reproducibility)

In terms of intra- and inter-laboratory reproducibility, 10 substances (two strongly active positives, four moderately active positives, one weakly active positive, and three negatives) were tested three times in each of three laboratories. All tests were conducted using stock solutions provided by CERI (i.e., the full test method protocol was not evaluated). Furthermore, substances that posed potential problems in testing due to their physico-chemical characteristics (i.e., poor solubility) or because they were overtly cytotoxic were not tested. Thus, this is not an adequate evaluation of the intra- or inter-laboratory reproducibility of this test method. In its international evaluation of another ER TA test method, NICEATM/ICCVAM is proposing 12 substances to evaluate intralaboratory reproducibility in three labs (testing 3 times in each lab) and another 41 substances to be tested once in each of three labs to adequately evaluate interlaboratory reproducibility. These substances cover the range of anticipated agonist and antagonist responses, include a wide variety of chemical classes, and include substances with varied physico-chemical properties and cytotoxicity properties.

Also, in their interlaboratory evaluation, the reference substance, estradiol, was tested over its complete concentration response range. In contrast, for other substances, CERI tested estradiol at a single

concentration. The former is recommended by the ICCVAM International ED Expert Panel and by ICCVAM for all experiments.

Test Method Data Quality

Interlaboratory studies testing 10 substances were conducted using GLP guidelines, but none of the pre-validation studies were conducted in this manner. At the last OECD preliminary validation assessment panel teleconference, CERI representatives indicated that a data audit has been recently conducted on the prevalidation studies and stated that noncompliance with GLP guidelines had no impact on data quality. We recommend that a specific discussion regarding data quality and noncompliance be included in the CERI report.

Animal Welfare Considerations (Refinement, Reduction and Replacement)

Our evaluation of the validation report and supporting materials indicate that specific discussions on how the proposed test method will refine, reduce, or replace animal use if used in a battery of tests to detect potential endocrine disruptors were not provided.

Practical Considerations

We recommend the inclusion of considerations such as the cost and time required to conduct the assay and report results. Considering the concerns about “edging effects”, we also recommend expanding the discussion of necessary equipment and supplies, and the required level of training, expertise and demonstrated proficiency needed by study personnel.

Appendix 2

Subject: FW: Statistical approach for intra- and interlaboratory variability

----- Forwarded Message

From: <Hasemanjk@aol.com>

Date: Thu, 25 May 2006 10:56:50 -0400

To: "Deal, Frank H (NIH/NIEHS) [C]" <dealf@niehs.nih.gov>

Cc: "Tice, Raymond (NIH/NIEHS) [E]" <tice@niehs.nih.gov>, "Ceger, Patricia (NIH/NIEHS) [C]" <cegerp@niehs.nih.gov>, "Blackard, Brad (NIH/NIEHS) [C]" <blackard@niehs.nih.gov>, "Charles, Jeffrey (NIH/NIEHS) [C]" <CharlesJ2@niehs.nih.gov>

Conversation: Statistical approach for intra- and interlaboratory variability

Subject: Re: Statistical approach for intra- and interlaboratory variability

Frank-

I have examined Dr. Aoki's PowerPoint slides, and I believe I understand his concerns.

The examples used to illustrate his concerns involve data from four labs with three runs per lab. These 12 datapoints (logPC50s) are apparently each based on estimates from a Hill equation analysis. However, regardless of how the estimates are obtained, each of the logPC50s is an estimate and has an associated SE of the estimate. One of Dr. Aoki's objections is that these standard errors associated with the estimation process are typically ignored in the data evaluation process.

For example, the typical approach for computing the mean response for each lab is to simply average the three runs. Dr. Aoki prefers instead a weighted average approach that weights each estimate inversely with the associated variability (i.e., the less variable estimate gets weighted more heavily in the averaging process). In my opinion, this is a reasonable option, and I suspect that a statistical purist would likely prefer the weighted average approach to the unweighted average. However, it could also be argued that since each run was carried out under identical conditions, the runs should be given equal weight, regardless of variability.

Thus, I disagree with Dr. Aoki that it is 'naïve' and 'inappropriate' to work with unweighted means, which provide unbiased estimates of the underlying parameter and typically are similar to the weighted means in any case. For example, in one of Dr. Aoki's examples, the unweighted mean is -6.94; the weighted mean is -6.93. I suspect that this is typical of what would be found in practice, especially since there are 'validity check' safeguards built in that will minimize the likelihood that the underlying variability estimates will differ greatly from run to run. From a practical point of view, it is unlikely in our area of application that the choice of weighted vs. unweighted means will have any noticeable impact on the overall interpretation of a study.

I note also that in Dr. Aoki's Slide 6, the lab and run columns are mislabeled and should be reversed.

A second related concern of Dr. Aoki is the calculation of an SD. For example, the variation in response among the three runs at a given lab in theory represents two distinct sources of variability: (i) the variability associated with the estimation process itself; and (ii) the additional variability that might be due to factors that are different from run to run. The SD that is normally calculated does not distinguish between these two sources of variability, but Dr. Aoki feels that this distinction is important and that by

subtracting out (i) and focusing strictly on (ii), one obtains better 'estimates'.

Better estimates of what? I agree that his approach provides better estimates of Source of Variability (ii), but I would argue that the primary variability of interest is the actual observed variability among runs, which reflects both (i) and (ii). It should not matter if this variability is due entirely to the estimation process (as was the case in three of the four labs in his example) or if both (i) and (ii) contribute to this variability. The end result is what matters.

Similar comments apply when combining the lab means to produce an overall average. Once again, one could either use a weighted average (-7.15 in Dr. Aoki's example) or an unweighted average (-7.13). Generally, the two will agree very closely.

The variability observed among the lab means is due to a combination of three sources of variability: (i) and (ii) as noted above and (iii) additional variability introduced by factors that differ among labs. Here again, Dr. Aoki recommends 'subtracting out' (i) and (ii) to obtain a 'pure' estimate of (iii). I would once again argue that it is the overall variability that is important, regardless of the contribution of the three individual components.

Although weighted versus unweighted means will very likely have little or no impact on the final interpretation of a study, the same may not be true for an evaluation of variability. In Dr. Aoki's 'fake data' example, he concludes that the much better SDs are essentially all zero. What does this mean from the standpoint of assessing the reproducibility of the assay? I worry that a naive investigator may assume that this means that the assay is extremely reproducible (after all, it has zero SD's), but this may not be the case at all. It may simply mean that the variability associated with the estimation process is so great that it can totally account for the overall variability in response observed among runs and among labs. The magnitude of this variability may or may not be cause for concern, but I still would argue that quantifying the specific sources of the variability is not nearly as important as evaluating the magnitude of the resulting variability itself, as assessed in the 'traditional' (and not 'inappropriate') way.

Dr. Aoki states in Slide 15 that the statistical programs used to produce the Hill equation estimates of the logEC50 do not provide associated SE estimates, but I do not believe that this is the case. Doesn't Prism produce them routinely? If so, then this information can be used in the manner suggested by Dr. Aoki.

Importantly, in the final analysis, one must decide if the purpose of these studies is to refine our estimates of the various sources of variability that contribute to differences in response, or is it to determine whether or not an assay has acceptable reproducibility. Dr. Aoki's presentation focuses on the former, but in my opinion, the latter should be our goal. Thus, if I am trying to determine whether or not an assay is acceptably reproducible, I would want to focus on the observed variability in the actual EC50 estimates across and within labs regardless of the factors that contributed to the variability.

For example, suppose I observed a coefficient of variation of 50%, that in normal circumstances would be unacceptable. However, using Dr. Aoki's approach, it is not this variability that is important, but the relative contribution of the factors that produced it.

This high variability might be due to the estimation process, differences among runs, differences among labs, or a combination of these three factors. In my opinion, quantifying these sources of variability and determining which is the primary contributor should not be our focus. For example, one extreme possibility is that the Hill equation model fit is so poor (and the resulting SE's of the estimated EC50's so high) that Source of Variability (i) can account for essentially all the variability in response, and as a result all the better estimate SD's computed by Dr. Aoki for Sources of Variability (ii) and (iii) are close to zero. Would Dr. Aoki consider such an assay to have acceptable reproducibility since the estimated SD's are all

close to zero? I would not.

If assessing the individual components contributing to the overall variability is viewed as a critical matter, then you could carry out a nested ANOVA to examine quantitatively the relative effects of variability among labs and variability among runs within labs on the overall response (e.g., the logEC50).

I could find nothing in Dr. Aoki's presentation to suggest how his approach could be used in a real world setting to determine whether or not an assay had acceptable reproducibility. One exercise that would be of interest would be to take a real world example and assess whether or not the assay has acceptable reproducibility in the usual way (considering CV's, etc.), and then ask Dr. Aoki and his colleagues to take the same data and make a similar 'bottom line' judgment based on his more complex assessment of weighted means, extracting sources of variability, etc. I strongly suspect that the same conclusion will be reached after considerably more work.

As a general rule, if a new complex statistical procedure is proposed to replace a 'less rigorous' one, then it should be demonstrated empirically how the old method fails and the advantages of the new approach in terms of the goal of the study, which in this case is accessing whether or not the assay has acceptable reproducibility. Until this is done, and concrete examples can be presented demonstrating the superiority of this more complex data assessment process, I see no need to make major changes in what is currently done.

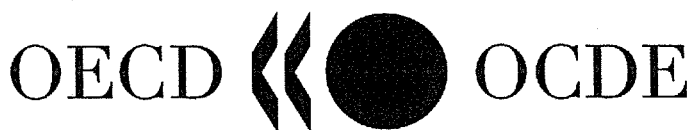
Regarding Appendix 2, I strongly agree with Dr. Aoki that it makes no sense to calculate a CV based on log-transformed data. Surely, no one is recommending this (are they?). If so, this should be abandoned, and I agree with Dr. Aoki that the measure of variability to use in this case is the SD, not the CV. I further agree with his assertion that 'In general, CV is a good measure of variation where SD of a variable increases (linearly) with the mean of the variable.'

Dr. Aoki then states that 'there seems to be no reason to believe that the SD increases with the mean'. It is unclear if he is referring to the SD associated with the log transformed data (in which case I agree with him) or the untransformed data (in which case I disagree).

For example, toxic compounds with very low EC50's may have three runs with estimated EC50 values of (e.g.) 0.01, 0.03, and 0.05, while a non-toxic compound may have EC50 values of 1000, 3000, and 5000. In such cases, the SD's of the EC50's are quite different, but the SD's of the log transformed data are identical. This is what generally happens in practice. Thus, in terms of the EC50 I would use CV; in terms of the logEC50 I would use SD. I suspect that Dr. Aoki would agree with this.

Joe Haseman
5-25-06

----- End of Forwarded Message



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

DIRECTION DE L'ENVIRONNEMENT
ENVIRONMENT DIRECTORATEDivision Environnement, santé et sécurité
Environment, Health and Safety Division

ENV/EHS/LM/jh/2006.21

Paris, 2006

To: National Co-ordinators of the Test Guidelines Programme

Cc: European Commission (DGs Environment, Enterprise, Sanco, Science, JRC, ECVAM)
 BIAC (including ACC, CEFIC, Croplife International, ECETOC, JCIA)
 EEB, ICAPO, ILSI Europe/North America, IPCS, TUAC
 Endocrine Disrupter Testing and Assessment Task Force
 Validation Management Group for Non Animal Testing
 ENV Counsellors to OECD Permanent Delegations (letter only)

Subject: Stably Transfected Transcriptional Activation (TA) Assay - Nomination of Peer Reviewers
Deadline: 30 September 2006

Dear National Coordinators,

Japan informed the Secretariat that the peer review of the validation of the above mentioned test method can start. This test method was presented and discussed by the VMG-non animal and an informal group chaired by Miriam Jabobs (OECD Secretariat) had several conference calls to prepare the peer review. The SPSF for this project is posted on the government protected website; I will send you soon a letter regarding several SPSFs received by the Secretariat.

The consultant who will manage the peer review has been identified and the Secretariat agreed to provide funding to contract the consultant; Japan already selected some candidate peer reviewers from Japan, France, United Kingdom and Germany.

In order to add names to the list of candidate peer reviewers, you are now invited to nominate experts for the peer review of the transfected TA assay validation, and provide corresponding CVs by **30 September 2006**. The expertise of the candidate peer-reviewers should be clearly mentioned in the CVs. Candidate peer-reviewers will be contacted directly by the consultant.

Best Regards,
 Laurence Musset
 Environment, Health and Safety Division

Task Description for the peer review manager

The task of the consultant will be to manage the peer review (PR) of the Japanese multi-laboratories validation study of a stably transfected estrogen receptor (ER)alpha mediated reporter gene assay developed by Japan. The consultant will:

1/ select a balanced international panel of reviewers with scientific and regulatory expertise amongst the experts proposed by Japan and nominated by the National Coordinators. The scientific reviewers should have expertise in at least one discipline, e.g., biology, toxicology, experimentation, validation or chemistry. Regulatory reviewers should, for example, have expertise in new chemicals, existing chemicals, pesticides or biocides. In addition, an observer should be selected to respond to questions about the validation (the Secretariat will provide a few names of people involved in the validation). If possible, another observer should be selected to provide responses regarding biostatistics;

2/ request and review declarations of interests from the panel members (ideally 5 to 10 members). The declarations of interests are not public documents. The consultant can select a peer reviewer even if he has declared interests, but the view of this peer reviewer should be balanced against his interests;

3/ ask questions to the peer review panel. The panel should be asked (i) whether the 8 validation criteria that are included in Guidance Document 34 are met, partially met or not met, and (ii) whether the reasons why a criteria is not met are acceptable; if the reasons are not acceptable, the peer review panel should provide recommendations on how to solve the problem. Of course, the consultant can add any other question(s), if he/she so wishes.

4/ collect written responses to the questions and identify issues which need to be further discussed by the panel;

5/ manage one or a few conference call(s) to try and come to an agreement (to facilitate the work of the consultant(s), the Secretariat can organize the conference calls but it will not participate);

6/ draft a PR panel report and circulate it for comments from the PR panel;

7/ provide Japan and the OECD Secretariat with (i) a peer review panel report (agreed by the PR panel), consisting of a general summary and a summary of PR panel responses to individual charge questions, (ii) the responses of each peer reviewer to the questions, and (iii) the names of the reviewers and their declaration of interests, at the latest by 31 March 2007.

Remarks

The consultants should be an independent manager of the process. He/She is not requested to provide technical/scientific input.

The new or updated Test Guidelines are expected to have been validated according to OECD criteria of validation, as described in the Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment (GD N° 34).

The peer review package will include the validation reports. In order to ensure transparency, the draft test guideline will be posted on the public Website, under a new heading "Peer Review Package".

The Secretariat can provide a document on how the 8 criteria of GD 34 have been addressed; this would be a useful document for the PR panel.

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の 編集者名	書 籍 名	出版社名	出版地	出版年	ページ
該当なし							

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
大野泰雄	日本薬理学会の奨める動物実験- 苦痛の評価と軽減- 「はじめに」および日本薬理学会の新動物実験指針	日本薬理学雑誌	129	5-9	2007
小島肇夫	動物実験代替に関する最近の動向	化粧品技術者会誌	40 (4)	263-268	2007
小島肇夫	JaCVAMの設立と使命	日皮協ジャーナル	57	129	2007

日本薬理学会の奨める動物実験 — 苦痛の評価と軽減 — 「はじめに」および日本薬理学会の新動物実験指針

大野 泰雄

要約：動物実験に対する社会の関心の高まりに伴い、平成 17 年 6 月に「動物の愛護及び管理に関する法律」が改正され、動物実験における 3R の原則を組み込まれた。また、文部科学省等の関連指針が改訂された。これらを考慮し、日本薬理学会は倫理的な動物実験を推進するために動物実験指針を作成した。本稿ではこれを紹介した。

最近の薬理学では細胞株等を用いる in vitro 研究が多くなっている。しかし、依然として、動物実験や実験動物から採取した試料を用いた研究は薬理学研究に不可欠である（表 1）。しかし、動物福祉や権利に対する社会の関心が高まり、動物実験への反対運動もたびたび報道されている。一方、現在の科学研究には多額の費用が必要であり、公的な資金なしに研究を進めることはできない。日本薬理学会はこのような状況に

表 1 第 77 回日本薬理学会でのポスター発表で用いられた試験系

試験系の種類	例数
In vivo 実験	185
薬物等で処理した動物から組織試料を採取して研究	32
動物から摘出した試料を用いて研究	118
	(335 69.6%)
In vitro 研究（細胞株等を用いた研究）	95
屠殺場から入手した試料を用いた研究	17
ヒト試料を用いた研究	19
アフリカツメガエル卵母細胞を用いた研究	7
	(138 28.7%)
その他（臨床試験、情報研究等）	8
	(8 1.7%)
合計	481

(2005 年 3 月 8 日および 9 日に発表されたものについての調査結果)

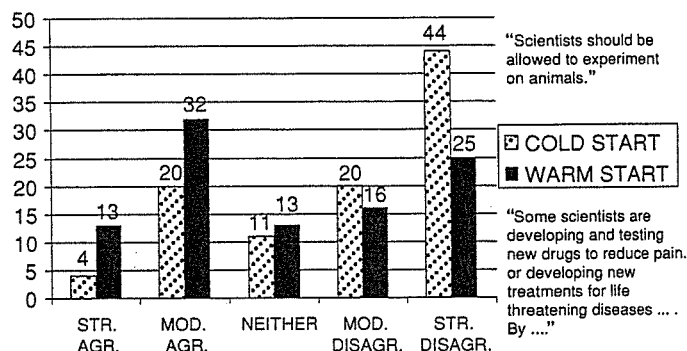


図 1 動物実験への一般人の反応

(New Scientist, 22 May, 1999, pp. 26-31, 1999 年英国での調査結果)

Cold Start: "Scientists should be allowed to experiment on animals." と説明した場合の反応

Warm Start: "Some scientists are developing and testing new drugs to reduce pain or developing new treatments for life-threatening diseases.... By...." と説明した場合の反応

STR.AGR: Strong Agreement

MOD.AGR.: Moderate Agreement, NEITHER: Neither

MOD.DISAGR.: Moderate Disagreement

STR.DISAGR: Strong Disagreement

適切に対応しなければならない。即ち、薬理学研究における動物実験の意義を社会に示し、科学的に必要なかつ倫理的に妥当な実験を行うことにより、我々の研究への社会の支持を得ることが不可欠である。図 1 に示したように、動物実験の必要性についての説明が十分になされることにより、科学的に必要な動物実験に賛同する者が確実に増加する。

日本薬理学会企画教育委員会では平成 16 年度より谷山絃太郎委員長の基、著者と昭和大学医学部の安原一教授および東京大学農学部尾崎博教授からなるワーキンググループ (WG) が組織された。WG では、

キーワード：動物実験、薬理学会動物実験指針

国立医薬品食品衛生研究所 (〒158-8501 東京都世田谷区上用賀 1-18-1)

e-mail: ohno@nihs.go.jp 原稿受領日：2006 年 10 月 2 日、会誌編集委員会依頼原稿

Title: Animal experiments recommended by the Japanese Pharmacological Society (evaluation and decrease of pain): introduction and new guidelines for animal experiments for the Japanese Pharmacological Society. Author: Yasuo Ohno

適宜、企画教育委員会の意見を求めながら、Bologna宣言（表2）で代表されるような動物実験に関する外国の状況や国内外の指針等を参考に、学会の動物実験指針の見直しを進めてきた。

一方、平成17年6月動物実験に関する3Rの原則が「動物の愛護及び管理に関する法律」（動愛法）に組み込まれ、薬理実験においても、研究や教育、生産などに使用される動物を用いる方法を他のものに置き換え、当初の目的を達すること（replacement）、特定の量と質を有するデータを得るために使用する動物を必要最小限にすること、また、同じ数の動物からより多くの情報をうること（reduction）、また、避けられない動物実験にあっては痛みや苦痛、および不快感を最低限にし、動物の福祉を向上させること（refinement）、が法的に義務づけられた（表3）。

更に、文部科学省、厚生労働省、および環境省ではそれぞれ所管する分野を対象に動物実験指針を作成し、平成18年4月-6月に通知した（環境省：「実験動物の飼育および保管並びに苦痛の軽減に関する基準（平成18年4月28日、環境省告示第88号）」、文部科学省：「研究機関等における動物実験等の実施に関する基本指針（平成18年6月1日、文部科学省告示第71号）」、厚生労働省：「厚生労働省の所管する動物実験等の実施に関する基本指針（平成18年6月1日、厚生労働省通知 科発0601002号）」が示された。また、日本学術会議も動物実験に関する詳細指針を作成し、通知した（平成18年6月1日、「動物実験の適正な実施に向けたガイドライン」）。日本薬理学会ではこれらを踏まえて、従来の動物実験指針を改訂した。平成18年3月の薬理学会総会では、平成17年12月13日の理

表2 ポロニア宣言

動物実験の削減、純化、および置き換え 代替法および実験動物法
<p>第三回生命科学における代替法と動物使用に関する世界会議 において採択（1999, 8,31 イタリア, ポロニア）</p> <p>Russell and Burch の「三つの R」、即ち削減（Reduction）、純化（Refinement）、および置き換え（Replacement）の源は 1954 年に開始された動物福祉のための大学連合（University Federation of Animal Welfare : UFAW）の活動にある。これが 1959 年に W.M.S. Russell and R.L. Burch (1) による「人道的な実験技術の原則（The Principles of Humane Experimental Technique）」の公刊につながっている。1978 年に David Smyth が代替法を三つの R として定義して使用した (2)。</p> <p>著書の中で、Russell と Burch は「科学における最も偉大な業績は常に最も人道的であり、かつ最も美的に引きつけるものであり、最も成功した時には科学の枢要である美しさと優雅さを感じさせるものである」と述べている。彼らは以下のように定義している。</p> <p>代替法における削減（Reduction Alternatives）とは科学的手法においてより少ない動物から同等の情報を得るための方法、あるいは同じ数の動物からより多くの情報を得るための方法である。</p> <p>代替法における純化（Refinement Alternatives）とは痛みや苦痛、および不快感を弱めたり、最少限にし、動物の福祉を向上させるものである。</p> <p>代替法における置き換え（Replacement Alternatives）とは動物を用いた実験や他の科学的な手段を用いずに当初の目的を達成するものである。</p> <p>1. Russell WMS & Burch RL (1959). The Principles of Humane Experimental Technique. p.238. London: Methuen. 2. Smyth D (1978). Alternatives to Animal Experiments. p. 218. London: Scholar Press.</p> <p>第三回生命科学における代替法と動物使用に関する世界会議の参加者は 1959 年に Russell & Burch により提起された原則を承認するとともに再確認するものである。人道的な科学とは善なる科学のための前提であり、かつ実験動物手法に関しては、三つの R を強力に推進し、適用することにより達成されるものである。</p> <p>三つの R は、全ての種類の科学、経済および人道主義的な便益を得るに際しての統一概念として、挑戦目標として、また、それらを得る機会として役立てるべきである。</p>

表3 改訂「動物の愛護及び管理に関する法律」抜粋

<p>第41条 動物を教育、試験研究又は生物学的製剤の製造の用その他の科学上の利用に供する場合には、科学上の利用の目的を達することができる範囲において、できる限り動物を供する方法に代わり得るものを利用すること、できる限りその利用に供される動物の数を少なくすること等により動物を適切に利用することに配慮するものとする。</p> <p>2 科学上の利用に供する場合には、その利用の必要な限度において、できる限りその動物に苦痛を与えない方法によってしなければならない。</p> <p>3 動物が科学上の利用に供された後において回復の見込みのない状態に陥っている場合には、その科学上の利用に供した者は、直ちに、できる限り苦痛を与えない方法によってその動物を処分しなければならない。</p>
--

事会で承認を受けた案をもとに説明したが、今回、各省庁の通知を踏まえて、新指針を作成したので、ここに紹介する。旧指針との対象表を表4に示したが、大きなところは以下のとおり。

- 1) 薬理学研究における動物実験の意義を明示した。
- 2) 薬理学研究が社会に受け入れられるためには、科学的・倫理的に適正な動物実験を行う環境を醸成し、実施することが不可欠であることを明示した。
- 3) 「動物が命あるものであることにかんがみ、何人も動物をみだりに殺し、傷つけ、又は苦しめることのないようにするのみでなく、人と動物の共生に配慮しつつ、その習性を考慮して適正に取り扱うようにしなければならない」という動愛法第2条に示された基本原則を明記した。
- 4) 3Rの原則を明示した。
- 5) 日本薬理学会は本指針に反する研究の成果は本会の刊行する学術雑誌から排除することを明示した。
- 6) 動物実験を適切に行うために必要な研究機関の責任者や実験動物の専門家等の役割を明示した。
- 7) 動物実験委員会の役割を明示した。

なお、平成18年3月のシンポジウムではこの動物実験指針改訂案を示すとともに、適切な実験を行うための参考とするために、奈良先端科学技術大学院大学の佐藤公道先生（現：安田女子大学）に苦痛の薬理学と薬理実験法について、秋田大学医学部実験動物施設松田幸久先生に痛み・苦痛・安楽死の評価と基準について、昭和大学医学部第二薬理安原 一先生と倉田知光先生に動物実験倫理委員会での審査の状況について、最後に大阪大学医学部動物実験施設黒澤努先生に欧米で一般化してきている動物実験施設の第三者による査察について報告していただいた。

著者プロフィール

大野 泰雄（おおの やすお）

◇東京大学薬学系大学院博士過程修了，
国立医薬品食品衛生研究所副所長，第六
回国際動物実験代替法会議会長。
◇専門：薬物代謝と毒性，動物実験代替
法。◇薬剤師，日本トキシコロジー学会
認定トキシコロジスト。

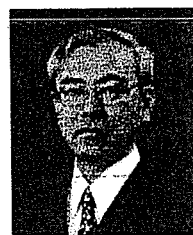


表4 日本薬理学会動物実験指針新旧比較

	旧 指 針	新 指 針
前 文		<p>生命科学の急速な発展と社会に与える影響の著しい拡大により，一般社会にとっても，生命科学研究がより身近なものになっている。また，研究は多額の公的資金によって支えられており，薬理学を含む生命科学研究の推進において社会の支持が不可欠の要素となっている。</p> <p>一方，動物を用いた研究は薬理学の発展に大きな役割を果たして来たとし，今後もその意義が失われることはないと考えられる。しかし，動物実験については，社会に様々な考え方が存在することも事実である。薬理学研究が社会に受け入れられるためには，科学的・倫理的に適正な動物実験を行う環境を醸成し，実施することが不可欠である。</p> <p>そこで，日本薬理学会では「動物実験ガイドラインの策定に関する勧告」（昭和55年11月5日 総学庶第1513号 日本学術会議会長）および「大学等における動物実験について（通知）」（昭和62年5月25日 文部省第141号 文部省学術国際局長）に定められている事項のほか，日本薬理学会員（以下，会員という）が動物実験を計画，実施する際に，遵守すべき基本的事項を定め，平成4年と13年に「動物実験に関する日本薬理学会指針」を学会員に通知し，科学的，倫理的観点から適正な実験動物の飼養と動物実験の実施に努めてきた。一方，動物福祉への社会の関心が更に高まり，平成17年6月15日に「動物の愛護及び管理に関する法律（動愛法）」が改正され，動物実験に関する3Rの原則*の尊重が盛り込まれた。また，平成18年4月28日に「実験動物の飼育及び保管並びに苦痛の軽減に関する基準（環境省告示第88号）」が，平成18年6月1日に「研究機関等における動物実験等の実施に関する基本指針（文部科学省告示 第71号）」，「厚生労働省の所管する動物実験等の実施に関する基本指針（厚生労働省通知 科発 0601002号）」，並びに日本学術会議から「動物実験の適正な実施に向けたガイドライン」が示されたことなど，わが国内外の動向も鑑み，指針を刷新することとした。</p> <p>日本薬理学会は本指針に従った動物実験が行われることを期待するとともに，これに反する研究の成果は本会の刊行する学術雑誌から排除する所存である。</p> <p>なお，遺伝子組換え動物に関しては，自然界への拡散を防止するため，「遺伝子組換え生物等の使用等の規制による生物の多様性の確保に関する法律（平成15年法律第97号）」ならびに「遺伝子組換え生物等の使用等の規制による生物の多様性の確保に関する法律施行規則（平成15年財務・文部科学・厚生労働・農林水産・経済産業・環境省令第1号）」が定められているが，これらの規制に関わる事項は対象外とした。</p> <p>*：充分な倫理的配慮を行った上，科学上の利用の目的を達する事が出来る範囲において，動物に与える苦痛を最小限にし（refinement），動物使用数を削減し（reduction），また，動物を用いない代替法がある場合にはそれを利用すること（replacement）。</p>

	旧 指 針	新 指 針
目 的	動物実験に関する日本薬理学会指針（以下、指針という）は、日本薬理学会員（以下、会員という）が動物実験を計画、実施する際に、「動物実験ガイドラインの策定に関する勧告」（昭和55年11月5日 総学席第1513号日本学術会議会長）および「大学等における動物実験について（通知）」（昭和62年5月25日 文学情第141号文部省学術国際局長）に定められている事項のほか、遵守すべき基本的事項を定め、科学的、倫理的観点から適正な実験動物の飼養と動物実験を実施することを目的とする。	この指針は、大学およびその他の研究機関において行われる薬理学研究のための動物実験を計画し、実施する際に遵守すべき事項を示すことにより、科学的にはもとより、動物福祉の観点からも適正な動物実験の実施を促すことを目的とする。
適 用 範 囲	この指針は、会員によって行なわれる実験動物を用いるすべての動物実験に適用されるものとする。	この指針は、会員によって行なわれる実験動物*を用いるすべての動物実験に適用されるものとする。 *：考慮の対象とする実験動物の範囲は基本的に生命を有する脊椎動物とその胚であるが、無脊椎動物が含まれることもある。また、これら以外も本指針を参考にする。
基 本 原 則		会員は「動物が命あるものであることにかんがみ、何人も動物をみだりに殺し、傷つけ、又は苦しめることのないようにするのみでなく、人と動物の共生に配慮しつつ、その習性を考慮して適正に取り扱うようにしなければならない」という動愛法第2条に示された基本原則を深く胸に刻み、ここに定められた事項を遵守するよう努めるとともに、動物実験に対する社会の動向や規制の移り変わりに留意し、常に適切な動物実験を実施するよう努めなくてはならない。 会員はヘルシンキ宣言のヒトを対象とする医学研究の倫理的原則（2002年追加）第12項に示された「研究に使用される動物の健康を維持し、または生育を助けるためにも配慮」や動愛法に示された動物実験に関する3Rの原則を尊重しなくてはならない。 薬理学研究を行う研究機関の責任者は、動物が適正に飼養され、適正な動物実験が行われるよう、施設を整備するとともに、研究機関毎の動物実験指針を策定し、研究者を教育しなければならない。また、動物実験委員会を設置し、研究機関内で行われる動物実験の法令や機関の定めた指針への適合性や科学的・倫理的妥当性を審査させるとともに、動物実験の実施結果の報告を受け、必要に応じて適正な動物実験実施のための改善措置をとらなければならない。 本指針の目的を達成することは、会員のみでの努力では困難である。研究機関の責任者や実験動物の専門家等の協力を得るよう努めなければならない。
実 験 者		動物実験を行おうとする研究者は動物実験を行うに際しての法令や規制・基準、倫理、麻酔法、鎮痛法、動物実験代替法についての教育、また、動物実験手技について訓練を受けていなければならない。
動 物 実 験 委 員 会		会員の属する研究機関においては、平成18年6月1日に示された文部科学省および厚生労働省の動物実験の実施に関する基本指針により、動物実験委員会を設置しなければならない。この委員会は動物実験が関係法令や機関の定めた指針に従い、科学的かつ倫理的に実施されるために動物実験計画を審査し、必要な助言を与え、また、適正な実施の監視を行う組織である。委員会は倫理的かつ科学的に妥当な動物実験を行う上で必要な知識と経験を有する実験動物の専門家、動物実験に関して優れた識見を有する者、その他必要と思われる者によって構成しなければならない。
動 物 実 験 場 所	動物実験は、適正に整備、管理された施設において、必要な設備のもとで行なわなければならない。	動物実験は、動物実験委員会が承認した、適正に整備、管理された施設において、必要な設備のもとで行なわなければならない。
実 験 動 物 飼 育 と 管 理	実験動物の取り扱いにあたっては、実験動物の生理、生態、習性ならびに飼育、管理方法に関する知識を十分にもたなければならない。	実験動物の入荷の際の検疫とその後の飼養については、そのための専門的な知識を有する動物管理責任者の協力を得て、適切な実験動物を確保すべきである。 動物実験の際の実験動物の取り扱いにあたっては、実験者自身も実験動物の生理、生態、習性ならびに飼育、管理方法に関する知識をもたなければならないが、それらの知識を十分にもつ専門家の助言を得ることも重要である。疾患モデル動物の作成や使用の場合においても同様である。

	旧 指 針	新 指 針
実験計画の立案	動物実験計画の立案にあたっては、その研究目的の達成に必要な最小限の実験にとどめ、適正な動物の選択および実験方法についての十分な配慮が必要である。また、適切な飼育環境（ケージの大きさ、収容動物数、温湿度、照明など）のもとに実験が実施できるよう実験計画を立案しなければならない。	動物実験計画の立案にあたっては、動物を用いないで、その研究目的を達成できる代替法の有無を考慮しなければならない。動物を用いる場合は、適正な動物種や系統を選択し、使用動物数と動物に与える苦痛を必要最小限にとどめるよう、実験方法についての十分な配慮が必要である。また、適切な飼育環境（ケージの大きさ、収容動物数、温湿度、照明など）のもとに実験が実施できるよう実験計画を立案しなければならない。 なお、実験計画は研究機関内の動物実験委員会による審査と承認を受けなければならない。
実験実施上の配慮	実験実施にあたっては、動物福祉の立場から、動物の不安や苦痛を、極力軽減するように務めなければならない。	動物実験は動物実験に熟達した者により、あるいはその指導のもとに行うべきである。また、動愛法および関連する規制・基準を遵守し、動物福祉の立場から、動物の不安や苦痛を、極力軽減するように努めなければならない。この際、国立大学法人動物実験施設協議会、NIH あるいは OECD の作成した安全性試験における人道的な指標に関するガイドラインが良い参考となる。 実験終了後の動物の取り扱いについては、「実験動物の飼養及び保管等に関する基準」（昭和 55 年 3 月 27 日 総理府告示第 6 号 平成 14 年 5 月 28 日一部改正）に従い、動物をすみやかに苦痛から解放するように努めなければならない。実験途中であっても、研究目的達成上不適切な強い苦痛が現れた場合には、動物をすみやかに苦痛から解放するように努めなければならない。 安楽死の方法については、国立大学法人動物実験施設協議会（2004）や日本獣医師会の解説（2000）を参照されたい。 動物実験および本指針遵守に関わる記録は適切に保管されなければならない。
実験終了後の処置	実験終了後の動物の取り扱いについては、「実験動物の飼養及び保管等に関する基準」（昭和 55 年 3 月 27 日 総理府告示第 6 号）に従い、動物をすみやかに苦痛から解放するように努めなければならない。	
安全管理上の配慮	物理的、化学的に注意を要する試料、または病原体を用いた動物実験を実施する場合には、施設管理者と協力し、一般留意事項、関係規則等を遵守して、安全の確保および環境汚染の防止のため十分な処置を講じなければならない。	物理的、化学的に注意を要する試料、または病原体を用いた動物実験を実施する場合には、施設管理者と協力し、一般留意事項、関係規則等を遵守して、安全の確保および環境汚染の防止のため十分な処置を講じなければならない。
その他	この指針に示されていない必要事項については、会員の所属する各研究機関における動物実験に関する諸規定、および「大学等における動物実験について（通知）」（昭和 62 年 5 月 25 日 文部省学術国際局長）を遵守するものとする。	この指針に示されていない必要事項については、会員の所属する研究機関における動物実験に関する諸規定、および「大学等における動物実験について（通知）」（昭和 62 年 5 月 25 日 文部省学術国際局長）を遵守するものとする。 日本学術会議第 7 部は 2004 年に「動物実験・施設の第三者評価機構の設置について」の提言を行っており、動物の飼育や管理、また、動物実験が適正に行われていることについて、第三者による認証を得ることも考慮しておく必要がある。 なお、動愛法の改正に伴い、文部科学省において動物実験指針の検討が始まった。それが完成した場合においては、必要に応じて本指針も改正しなければならない。
引用文献		国立大学法人動物実験施設協議会：動物実験処置の苦痛分類に関する解説、平成 16 年 6 月 4 日 (http://www.med.akita-u.ac.jp/~doubutu/kokudou/rinri/pain.pdf) NIH：Guide for the Care and Use of Laboratory Animals (http://www.nap.edu/readingroom/books/labrats/) OECD：Guidance document on the recognition, assessment and use of clinical signs as humane endpoints for experimental animals used in safety evaluation. Environmental Health and Safety Publications. Series on Testing and Assessment No. 19 (2000.11). 日本実験動物環境研究会：「実験動物の飼養及び保管等に関する基準」についての日本実験動物環境研究会改正案。実験動物と環境 Vol.12 (1), 71-74, 2004. 鈴木真, 黒澤努：日本獣医師会雑誌。一解説・報告—米国獣医師会：安楽死に関する研究報告 Vol.58(5)301-304, (6)357-359, (7)443-446, (8)521-524, (9)581-583, (10)649-651, (11)719-721, 2000.

J. Soc. Cosmet. Chem. Jpn.
総 説
40(4) 263 — 268 (2006)

動物実験代替に関する最近の動向*

小 島 肇 夫

国立医薬品食品衛生研究所 安全性生物試験研究センター
薬理部 新規試験法評価室**

動物実験における 3Rs (削減, 苦痛の軽減, 置換) の推進のため, EU および米国は, 1990 年代初頭, それぞれ代替法のバリデーションおよび評価を進めるため, European Center for the Validation of Alternative Methods (ECVAM) や Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) を設立した。これらの機関は, 試験法の行政的な受入れを目的に, 代替法の信頼性および適性をバリデーションや評価により確認することを目的としている。日本でも 2005 年に国立医薬品食品衛生研究所内に Japanese Center for the Validation of Alternative Methods (JaCVAM) が設立された。しかし, JaCVAM だけでことをなすには資金や人員もなく, 日本動物実験代替法学会や日本化粧品工業連合会などの団体の援助が必要である。JaCVAM が多くの支援のもとで, 日本における化粧品の安全性評価のための代替法研究のまとめ役として機能する日も近いと考えている。

1. 代替法とは

昨年 11 月, 国立医薬品食品衛生研究所 (National Institute of Health Sciences: NIHS) 安全性生物試験研究センター, 薬理部内に新規試験法評価室という部署が設立された。この部署の英語名は (Japanese Center for the Validation of Alternative Methods) である。通称を JaCVAM という。日本語と英語を比較された方から, 合っていないという意見をよく頂く。確かに英語を直訳したら, 日本代替法バリデーションセンターとでもなるのか。名前から受ける印象は, 組織構成も扱う範囲も別のように感じられるかもしれないが, Validation of Alternative Methods が主な仕事である。Alternative という言葉は英和中辞典 (研究社) によると代わりのという意味なので, replacement という言葉の置き換えと同義語と勘違いされている方は多く, 英語名では動物実験を置き換えるための方

法, すなわち *in vitro* 試験による replacement を目的とする部署でしょうか? と誤解される。しかし, OECD ガイドライン GD 34¹⁾によると, Alternative test とは以下のように定義されている。

Alternative test: A test that: reduces the numbers of animals required; refines procedures to lessen or eliminate pain or distress to animals, or enhance animal well-being; or uses non-sentient material in replacement for conscious living vertebrates

つまり, JaCVAM の主な使命は, 3Rs (Reduction, Refinement, Replacement) を強く意識し, *in vitro* 試験法の確立以外にも動物実験を削減したり, 苦痛やストレスを削減したり, 動物愛護の促進に関与するすべての試験, ガイドラインに関与したものでなければならぬと考えている。よって, 3Rs に関係した内容すべてに関与していくという姿勢が JaCVAM の仕事と考えている。

その考えでいけば, 化粧品の安全性に関する代替法の確立はその一部分でしかない。本年 6 月から施行された動物の愛護および管理に関する法律

* 2006. 7. 19 受理

** 〒158-8501 東京都世田谷区上用賀 1-18-1: 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

の一部改正²⁾を、関与するすべての動物実験を対象として浸透させていく立場にもある。REACH (化学物質の登録, 評価, および認可のシステム) に代表される化学物質の安全性評価³⁾, 日米 EU 医薬品規制調和国際会議 ICH (International Conference of Harmonization)⁴⁾でも, 動物福祉の面から考慮がなされている。これらの情報の供給を書籍やホームページで発信せねばならない。場合によっては動物実験を削減する代わりにヒトの臨床試験方法を整備するなど範囲に含まれるかもしれない。その守備範囲は大変広いと考えている。

動物実験を巡る世界の動向や新試験法の成立までの知識がなく, ここまで, 読んでいただけた方には敬意を表したい。時計の軸を戻しながら, JaCVAM 設立の経緯からまず説明していきたい。

2. JaCVAM 設立の経緯と活動

動物実験代替法のブーム発端は, 1980 年代に化粧品の誤使用を想定した眼刺激性試験 (ドレイズ試験) での化粧品の安全性評価において, ウサギの眼の使用が残酷というイメージが先行し, 愛護団体が運動を強めたことによる。この盛り上がりを受け, 欧米を中心にこの試験の代替法開発が進み, 多くの論文が発表された。代替可能と読み取れるこれらの論文から, 代替法が確立されたと錯覚させるような時代であった。しかし, 実際は再現性に乏しかったり, 被験物質が偏っていたりで思ったような良い試験法は少なかった。この反省から, 施設内および施設間評価による信頼性,

再現性を重視した試験法のバリデーション, 評価という検証システムが定着してきた^{5), 6)}。このシステムを Fig. -1 に示す⁶⁾。ある著名な毒性学者が新試験法を開発され, 特許を取られたとしても, その試験法がそのまま普及して申請資料などに使える, 企業化可能なものになる訳ではない。安全性確保のための行政的な受入れには, 複数施設の協力により, ブラインド化したある程度の被験物質を用いたバリデーションが必要であり, さらにその後, その分野の専門家による評価を経て, 行政的な受入れが決まる。このバリデーション, 評価を専門に行う機関として, 1990 年初頭に欧州では ECVAM (European Center for the Validation of Alternative Methods), 米国では ICCVAM (Inter-agency Coordinating Committee on the Validation of Alternative Methods) が設立された。

欧米では動物愛護団体の力が強く, 数も多い。その力は政治をも動かすこともあり, その結果 ECVAM や ICCVAM が設立されたと考えてもよい。また, 特に化粧品のよう嗜好性の高い商品の開発に動物実験を用いることへの風当たりは強く, 2003 年に採択された EU directive 7 次改正において, 2004 年から EU 域内における最終製品の動物実験禁止が適用され, 2009 年をもって, 最終製品および化粧品成分に対する動物実験禁止, および販売の禁止が適用される⁷⁾。ただし, 代替法の開発と歩調をそろえなければならず, 反復投与毒性, 生殖毒性試験, トキシコキネクスなどの試験は 2013 年が最終期限である。これに合わせるように, 欧米の化粧品会社は試験法の開発のため共同研究を行ったり, 国際会議や公的な機関に多くの資金を投入している現状である。もちろん, 国際化が進む昨今, 日本のメーカーも例外ではなく, 歩調を合わせる必要が生じている。

このような状況の中, 日本でも遅まきながら, バリデーションや評価の統括を専門に行う機関であり, かつ欧米の専門機関への窓口として昨年 JaCVAM が設立され, Fig. -2⁸⁾ に示すように代替法評価のための 3 機関が確立された。お断りしておくが, この 10 年間まったくのブランクが日本にあった訳ではない。

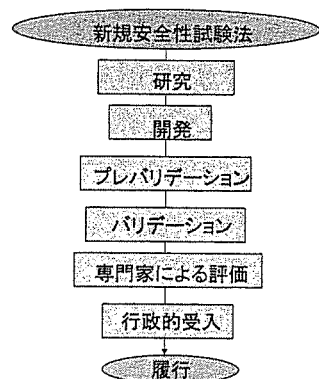


Fig. -1 Test method validation process.⁶⁾