European Commission

ESAC Opinion on the use of Performance Standards to evaluate test methods similar to a Validated Reference Method

ESAC Opinion No. 2016-06
of 24 June 2016

2016

EUR 28183 EN

Contact information
Name: João BARROSO
Address: Joint Research Centre, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy
E-mail: Joao.BARROSO@ec.europa.eu
Tel.: +39 0332 78 5329

JRC Science Hub
https://ec.europa.eu/jrc

JRC103708

EUR 28183 EN

PDF ISBN 978-92-79-63211-2 ISSN 1831-9424 doi:10.2787/581048

Luxembourg: Publications Office of the European Union, 2016

How to cite: EURL ECVAM Scientific Advisory Committee (2016). ESAC Opinion on the use of Performance Standards to evaluate test methods similar to a Validated Reference Method. ESAC Opinion No. 2016-06 of 24 June 2016; EUR 28183 EN; doi:10.2787/581048. Available at: http://publications.jrc.ec.europa.eu/repository/handle/JRC103708.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
JOINT RESEARCH CENTRE
Directorate F - Health, Consumers and Reference Materials
European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)

EURL ECVAM
SCIENTIFIC
ADVISORY
COMMITTEE
(ESAC)

ESAC OPINION

on the

Use of Performance Standards to Evaluate Test Methods Similar to a Validated Reference Method

ESAC Opinion No.	2016-06
Relevant ESAC Request No.	Not Applicable
Date of Opinion	24/06/2016

Table of contents

Abstract

ESAC, the EURL ECVAM Scientific Advisory Committee, advises EURL ECVAM on scientific issues. Its main role is to conduct independent peer review of validation studies of alternative test methods and to assess their scientific validity for a given purpose. The committee reviews the appropriateness of study design and management, the quality of results obtained and the plausibility of the conclusions drawn. ESAC peer reviews are formally initiated with a EURL ECVAM Request for ESAC Advice, which provides the necessary background for the peer-review and establishes its objectives, timelines and the questions to be addressed. The peer review is normally prepared by specialised ESAC Working Groups. These are typically composed of ESAC members and other external experts relevant to the test method under review. These experts may be nominated by ESAC, EURL ECVAM and partner organisations within the International Cooperation on Alternative Test Methods (ICATM). ESAC ultimately decides on the composition of these Working Groups. ESAC's advice to EURL ECVAM is formally provided as 'ESAC Opinions' and 'Working Group Reports' at the end of the peer review. ESAC may also issue Opinions on other scientific issues of relevance to the work and mission of EURL ECVAM but not directly related to a specific alternative test method.

The ESAC Opinion expressed in this report relates to the use of Performance Standards to evaluate test methods similar to a Validated Reference Method.

Ispra, 24 June 2016

ESAC Opinion

At the 42nd meeting of the EURL ECVAM Scientific Advisory Committee (ESAC) (Annex 1), two *in vitro* toxicology methods were evaluated according to the published OECD Performance Standards (OECD, 2015a, b). In the discussion several issues came to light which are generic in nature and concern the general approach to Performance Standards.

The "Performance Standards" approach has four main objectives:

1. To allow competition between developers of test methods.
2. To encourage the development of "me too/generic" test methods.
3. To ensure that these "me too/ generic" methods, perform equally as well as the original, fully validated method.
4. To allow the evaluation of the performance of these "me too" methods to be more efficient and less costly than a full validation.

ESAC fully supports these aims which it believes can be achieved, as intended, but requires a different approach to evaluating the data, viz.

While power calculations have identified that full validation studies require more than 50 chemicals to be evaluated in order to calculate sensitivity and specificity and hence accuracy with satisfactory precision, these cannot be meaningfully calculated with e.g. 10 negative and 10 positive reference chemicals as proposed in the OECD Series on Testing and Assessment No. 220 (OECD, 2015b). The 95 %-Confidence Interval for a specificity of 8/10 = 80 % observed in the "Performance Standards" approach ranges from 44 % to 97 %. From this it can be seen that there is insufficient statistical power to evaluate a new test method's performance using this measure.

The same is true of comparing within and between laboratory reproducibility where the sample size is equally small.

As the reference chemicals proposed in Performance Standards to evaluate "me too/ generic" test methods will have been used in the validation of the reference method(s), it seems more useful to assess equivalence of a new method with the reference method(s), side by side, on the basis of the predictions for the individual reference chemicals instead of comparing sensitivity, specificity and accuracy to prespecified threshold values. It is, e.g. possible that two methods meet the performance standards of 80 % for sensitivity, specificity and accuracy, but disagree on 4 of the 20 reference chemicals (this would correspond to a Fleiss Kappa value of 0.6, indicating only moderate agreement between the methods).

It should be possible in the Performance Standards instead to specify the maximum number of different results allowed if the method is considered to be equivalent. Scientific judgment should be used in examining and concluding on these differences. For example, one method may have discordant runs with a majority of runs positive for a given endpoint and for the same test chemical the other method has discordant runs with the majority negative. The fact that the result was based on discordant runs in both cases might suggest that there was no real difference between the methods.

More flexibility in the application of the Performance Standards would be helpful in obtaining more meaningful results, for example:

Provision should be made to exclude from the maximum permitted number of five independent tests per reference chemical those tests included in non-qualified runs that occur due to failure of the Positive or Negative Controls; and/or due to technical issues that are legitimately identified whilst conducting the test run (e.g. obvious chemical or microbiological contamination), provided that the testing is then stopped and the run abandoned at that point. If such a provision is included in the Performance Standards, then a maximum permitted number of non-qualified runs should also be defined.

However, in some situations, less flexibility is warranted. In order to have confidence in the conclusions of such evaluations, ESAC believes it is important to have more rigorous enforcement with respect to the use of different chemical sets to develop and optimise the test method, for training and proficiency testing, and for ring-trial reference chemicals.

References

- OECD (2015a). Performance Standards for assessment of proposed similar or modified *in vitro* skin sensitisation ARE-Nrf2 Luciferase test methods. Series on Testing and Assessment No. 213.
- OECD (2015b). Performance Standards for the assessment of proposed similar or modified *in vitro* Reconstructed human Epidermis (RhE) test methods for skin irritation testing as described in TG 439 (Intended for the developers of new or modified similar test methods). Series on Testing and Assessment No. 220.

Annex 1

COMPOSITION OF ESAC

Composition of ESAC

EURL ECVAM Scientific Advisory Committee (ESAC)

- Dr. Neil CARMICHAEL (ESAC Chair)
- Prof. Jürgen BORLAK
- Dr. Harvey CLEWELL
- Prof. Lucio G. COSTA
- Dr. Kristina KEJLOVÁ
- Prof. David John KIRKLAND
- Prof. Annette KOPP-SCHNEIDER
- Dr. Renate KRÄTKE
- Prof. Claus-Michael LEHR
- Dr. José Maria NAVAS
- Prof. Aldert PIERSMA
- Dr. Jonathan RICHMOND
- Dr. Erwin L. ROGGEN
- Dr. Dorothea SESARDIC

EURL ECVAM (Secretariat)

- Dr. João BARROSO (ESAC Coordinator)
- Prof. Maurice WHELAN (Head of Unit)

LA-NA-28183-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2787/581048

ISBN 978-92-79-63211-2