Regulatory Toxicology and Pharmacology 65 (2013) 278–285

Contents lists available at SciVerse ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Performance standards and alternative assays: Practical insights from skin sensitization ☆

Susanne N. Kolle [a,*], David A. Basketter [b], Silvia Casati [c], William S. Stokes [d], Judy Strickland [e], Bennard van Ravenzwaay [a], Hans-Werner Vohr [f], Robert Landsiedel [a]

[a] BASF SE, Experimental Toxicology and Ecology, Ludwigshafen, Germany
[b] DABMEB Consultancy Ltd., Sharnbrook, United Kingdom
[c] EURL ECVAM, JRC Ispra, Italy
[d] National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA
[e] Integrated Laboratory Systems, Inc., Research Triangle Park, NC, USA
[f] Bayer HealthCare AG, Bayer Pharma, Wuppertal, Germany

ARTICLE INFO

Article history:
Received 21 October 2012
Available online 29 December 2012

Keywords:
Performance standards
Validation
Local lymph node assay
Lymph node cell counts
Skin sensitization
Alternatives

ABSTRACT

To encourage the development and validation of alternative toxicity test methods, the effort required for validation of test methods proposed for regulatory purposes should be minimized. Performance standards (PS) facilitate efficient validation by requiring limited testing. Based on the validated method, PS define accuracy and reliability values that must be met by the new similar test method. The OECD adopted internationally harmonized PS for evaluating new endpoint versions of the local lymph node assay (LLNA). However, in the process of evaluating a lymph node cell count alternative (LNCC), simultaneous conduct of the regulatory LLNA showed that this standard test may not always perform in perfect accord with its own PS. The LNCC results were similar to the concurrent LLNA. Discrepancies between PS, LLNA and LNCC were largely associated with "borderline" substances and the variability of both endpoints. Two key lessons were learned: firstly, the understandable focus on substances close to the hazard classification borderline are more likely to emphasise issues of biological variability, which should be taken into account during the evaluation of results; secondly, variability in the results for the standard assay should be considered when selecting reference chemicals for PS.

1. Introduction

The validation of alternative toxicology tests represents a challenging, but important, requirement to demonstrate the relevance and reliability of their predictive value in respect of a stated purpose. To this end, the requirements of validation have been extensively published and debated. In particular, experience of previous formal validation activities, such as that for the local lymph node assay (LLNA[1]), the first ever completed (ICCVAM, 1999; Dean et al., 2001), has been used to help refine both the validation process and its detailed requirements (Sailstad et al., 2001; ICCVAM, 2003). However, the impetus for the development of *in vitro* alternatives has increased further the pressure to accelerate the validation process, there being no doubt that the said process is viewed by some as potentially onerous and by some as obstructive to timely progress (Basketter et al., 2010). One element of the response to these matters has been the proposal to adopt performance standards (PS) (ECVAM, 2008; ICCVAM, 2003, 2009; OECD 2010a). These are seen as a way to establish a minimum set of requirements for an alternative assay and, further, as a means of establishing a "level playing field" for those engaged in the development of alternatives (Stokes et al., 2006; Stokes and Wind, 2010; Wind and Stokes, 2010).

In the specific case that forms the subject matter of the present paper, PS had been established in a coordinated manner between different validation authorities. The aim of these PS was to facilitate the development and acceptance of minor variations to the standard LLNA (OECD, 2002, 2010a), such as replacing the use of tritiated thymidine with a non-radioactive measurement. The almost immediate positive contribution of these PS was seen in the rapid assessment of two LLNA variants (reviewed in Basketter et al., 2008; ICCVAM, 2010a, 2010b), whose success triggered the

☆ This article may be the work product of an employee or group of employees of the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), or other organizations, however, the statements, opinions or conclusions contained therein do not necessarily represent the statements, opinions, or conclusions of NIEHS, NIH, United States government, or other organizations.

* Corresponding author. Address: BASF SE, Experimental Toxicology and Ecology, Z570, 67056 Ludwigshafen, Germany.
E-mail address: susanne.kolle@basf.com (S.N. Kolle).

[1] *Abbreviations*: ACD, allergic contact dermatitis; CV, coefficient of variation; EC, estimated concentration; ECVAM, European Centre for the Validation of Alternative Methods; ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; LNCC, local lymph node cell count; LLNA, local lymph node assay, OECD, Organisation for Economic Co-operation and Development, PS, performance standards.

0273-2300/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yrtph.2012.12.006

adoption of two new OECD test guidelines (OECD, 2010b, 2010c). Neither of these variants to the standard LLNA used cell counting in the draining lymph nodes, an approach which had been suggested a number of years earlier and for which the extensive experience of one company has recently been published (Basketter et al., 2012; Kolle et al., 2012). Within this body of work, the LLNA PS were used to evaluate a protocol variation which employed cell counts as an alternative readout (LNCC); a concurrent LLNA was conducted. The outcome of this work and the light that it sheds on both the LLNA and on the development and interpretation of PS offers a valuable learning opportunity that may have general value in the world of toxicology and the evaluation of alternatives.

2. Materials and methods

Most of the LNCC and LLNA results referred to in this paper have been published previously (Basketter et al., 2012; Kolle et al., 2012). Importantly, LLNA and LNCC data reported by Basketter et al. and Kolle et al. were obtained from the same animals. Each study was carried out under GLP conditions, according to OECD, European Commission, and US EPA test guidelines (OECD, 2002, 2010a; EU, 2004; US EPA, 2003), and published protocols and modifications of the Murine Local Lymph Node Assay (LLNA) (Gamer et al., 2008; Basketter et al., 2012). In addition to the previously published data, LLNA and LNCC results for four PS were conducted as confirmatory studies in a contract research laboratory in a blinded manner. Further 13 PS were evaluated in the LNCC at Bayer HealthCare AG were included in the current analysis.

Details of the LLNA and LNCC methods were described fully in Basketter et al. and Kolle et al. Briefly, 3 groups of 5 mice were treated with the 3 concentrations of the test substances and one group was treated with vehicle by applying them to the dorsal part of both ears (25 μL per ear for 3 consecutive days at the same site). Three days after the last application, the mice were injected intravenously into the tail vein with 20 μCi of ^{3}H-thymidine in 250 μL of sterile saline for the LLNA. Mice were sacrificed 5 h later. Left and right auricular lymph nodes were dissected, weighed and prepared for determination of cell count and ^{3}H-thymidine incorporation. The responses to test substances exposures were characterized by lymph node cell count/lymph node pair (LNCC), ^{3}H-thymidine incorporation into the lymph node cells (LLNA), and ear weight. The skin sensitizing potential of a test substance is indicated by an increase in the stimulation index (SI) of the LNCC by a factor of $\geqslant$1.5 and/or of ^{3}H-thymidine incorporation by a factor of $\geqslant$3 compared to the concurrent vehicle control group. Where applicable, the EC (estimated concentration) leading to the respective threshold SI values, 3 for the LLNA and 1.5 for the LNCC, were calculated by linear or semi-logarithmic regression between the data points directly below and above the SI if possible or using the two nearest points below or above the SI (Basketter et al., 1999a, 1999b; Ryan et al., 2007).

For ease of description, the key results in a simplified form are detailed in Table 1. Additional LLNA data are appropriately referenced. Further analysis has been undertaken on substances that had discordant results between LLNA experiments and/or between the LLNA and the LNCC. To complement this, the scientific and clinical literature have been examined to provide information on the PS substances to help address the point of their relative potency as skin sensitizers in humans and in the LLNA. This is the most appropriate way to determine whether they are borderline substances in terms of classification, which would have a consequent impact on the reproducibility of test predictions.

3. Results

The LLNA PS comprises a total of 22 substances (ICCVAM, 2009; OECD, 2010a). The expected result for each of these is detailed in the last column of Table 1. Discrepancies between the LLNA PS and the concurrent LLNA (shown in column 2 of Table 1) arose for 5 of these substances: chlorobenzene, methyl salicylate, methyl methacrylate, nickel chloride and salicylic acid. Examination of the data supporting the PS indicates that for 3 of these substances, the

Table 1
Overall test results of testing of PS substances.

Substance	LNCC[a]		LLNA[b]		PS LLNA[c]	
	Result	EC1.5 (%)	Result	EC3 (%)	Result	EC3 (%)[d]
MCI/MI	+	0.011	+	0.01	+	0.009
Dinitrochlorobenzene	+	<0.025	+	<0.025	+	0.049
Phenylenediamine	+	0.1	+	0.1	+	0.11
Cobalt chloride	+	<0.25	+	<0.25	+	0.6
Isoeugenol	+	2.7	+	2.2	+	1.5
Mercaptobenzothiazole	–	Neg	+	4.6	+	1.7
Citral	+	9.2	+	12.6	+	9.2
Eugenol	+	8.7	+	4.6	+	10.1
Hexylcinnamal	+	16.2	+	9.2	+	9.7
Phenyl benzoate	+	4.0	+	8.9	+	13.6
Cinnamic alcohol	+	26.0	+	25.2	+	21
Imidazolidinyl urea	+	21.5	+	15.9	+	24
Methyl methacrylate	–	Neg	–	Neg	+	90
Chlorobenzene	+	79.0	+	45.6	–	Neg[e]
Isopropanol	–	Neg	–	Neg	–	Neg
Lactic acid	–	Neg	–	Neg	–	Neg
Methyl salicylate	+	49.0	+	32.8	–	Neg
Salicylic acid	+	15.8	+	8.0	–	Neg
Sodium lauryl sulfate	+	1.6	+	2.9	+	8.1
Ethyleneglycol dimethacrylate	+	38.0	+	45.1	+	28
Xylene	+	28.2	+	39.1	+	95.8
Nickel chloride	+	3.6	+	3.5	–	Neg

[a] Local lymph node assay based on cell counts, as published by Basketter et al. (2012).
[b] Concurrent LLNA as published by Basketter et al. (2012).
[c] Expected LLNA result contained in the PS documentation (ICCVAM 2009; OECD 2010a) (note: substances presented in the order of that documentation).
[d] Values taken from the PS documentation (ICCVAM 2009; OECD 2010a).
[e] Neg reflects the fact that potency values cannot be derived where the test result is negative.