National Institute of Environmental Health Sciences
Your Environment. Your Health.

Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers

National Institutes of Health
U.S. Department of Health and Human Services

Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
National Institute of Environmental Health Sciences
National Institutes of Health
Department of Health and Human Services

August 2025

Foreword

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) is an office of the Division of Translational Toxicology within the National Institute of Environmental Health Sciences. NICEATM focuses on the development and evaluation of alternatives to animal use for chemical safety testing. It was established by the ICCVAM Authorization Act of 2000 (42 U.S.C. 285 *l*-3) to provide support to the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). NICEATM and ICCVAM work collaboratively to evaluate new and improved testing approaches applicable to the needs of U.S. federal agencies.

NICEATM provides technical and scientific support for ICCVAM and ICCVAM workgroup activities, peer review panels, expert panels, workshops, and validation efforts.

In addition to providing support for ICCVAM, NICEATM:

- Supports National Toxicology Program activities, especially those contributing to the U.S. government's interagency Tox21 initiative.
- Conducts analyses and evaluations and coordinates independent validation studies on novel and high-priority alternative testing approaches.
- Provides information to test method developers, regulators, and regulated industry through its website and workshops on topics of interest.

NICEATM publishes reports of its test method development and evaluation activities in the scientific literature. NICEATM also issues reports of ICCVAM test method evaluations and other communications and makes these available on the National Toxicology Program (NTP) website, where they are available free of charge. Data from these studies are available via both NICEATM's Integrated Chemical Environment and NTP's Chemical Effects in Biological Systems database.

For questions about NICEATM and ICCVAM reports and studies, please contact NICEATM.

Preface

Regulators worldwide require manufacturers to test pesticides, cosmetics, household cleaners, and other chemical products prior to marketing to identify skin sensitizers, or substances with the ability to cause allergic contact dermatitis. There is a widespread international need for non-animal test methods that can identify these substances.

The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is charged with ensuring that new and revised test methods are validated to meet the needs of U.S. federal agencies. To advance availability and acceptance of non-animal methods to identify potential skin sensitizers, the National Institute for Occupational Safety and Health submitted a nomination to National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which provides support to ICCVAM, for an in chemico test method for allergic contact dermatitis hazard testing: the electrophilic allergen screening assay (EASA).

The nomination requested that NICEATM and ICCVAM evaluate the EASA as a screening assay for identifying contact allergens. A preliminary evaluation by NICEATM concluded that the EASA should be further evaluated for its usefulness and limitations for allergic contact dermatitis hazard classification. The preliminary evaluation stipulated that optimization and standardization of the EASA test method protocol was required prior to conducting an inter-laboratory validation study to characterize the reliability and relevance of the test method. Petersen et al. (2022a) describes standardization of the EASA and its modification to a 96-well format to increase throughput and accessibility of the assay.

This report summarizes the key findings of a validation study for the EASA conducted by four laboratories, representing a collaboration among five ICCVAM member agencies. The study assessed the EASA's within- and between-laboratory reproducibility and its performance against in vivo reference data. The report also discusses the purpose and scientific rationale of the EASA, describes the conduct of the assay and problems encountered during the validation process, and addresses the EASA's predictive capacity and applicability domain. The information from this study may support future consideration of the EASA for addition to existing accepted test guidelines for similar methods, and whether the EASA could be used as part of an integrated approach for testing and assessment and/or a defined approach to discriminate between skin sensitizers and nonsensitizers for hazard classification and labeling and potency categorization.

NICEATM's activities are guided in part by the "Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States" issued by ICCVAM in 2018. One objective articulated in the Strategic Roadmap was that ICCVAM agencies would utilize public-private partnerships to promote cross-sector communication and cooperation. An implementation plan developed for the Strategic Roadmap stated that NICEATM, ICCVAM, and collaborators advances the use of integrated approaches to testing and assessment and defined approaches to enable prediction of skin and eye irritation hazard. The project described in this report addresses both these objectives.

Table of Contents

Figures

Tables

Abbreviations Used

ACD	Allergic contact dermatitis
ACN	Acetonitrile
ADRA	Amino acid derivative reactivity assay
APHC	U.S. Army Public Health Center (now Defense Center for Public Health – Aberdeen)
AOP	Adverse outcome pathway
BLR	Between-laboratory reproducibility
BRT	Burleson Research Technologies, Inc.
CDER	Center for Drug Evaluation and Research
CDRH	Center for Devices and Radiological Health
COV	Coefficient of variation
CPSC	U.S. Consumer Product Safety Commission
DA	Defined approach
DCPH-A	Defense Center for Public Health – Aberdeen
DoD	Department of Defense
DPRA	Direct peptide reactivity assay
DTT	Division of Translational Toxicology
EASA	Electrophilic allergen screening assay
EC3	Dose (weight %) that gives a stimulation index of 3, positive threshold
EPA	U.S. Environmental Protection Agency
FDA	U.S. Food and Drug Administration
IC_{50}	Half-maximal inhibitory concentration
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
IQR	Interquartile range
kDPRA	Kinetic DPRA
KE	Key event
LLNA	Local lymph node assay
MIE	Molecular initiating event
NBT	4-Nitrobenzenethiol
NC	Negative control
NICEATM	National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods
NIEHS	National Institute of Environmental Health Sciences

NIST	National Institute of Standards and Technology
NTP	National Toxicology Program
OECD	Organisation for Economic Cooperation and Development
PC	Positive control
PDA	Pyridoxylamine
PS	Performance standards
QC	Quality control
QT	Qualified test
SS	Solvent system (1:1 acetonitrile:phosphate buffer)
TC	Test chemical
TG	Test guideline
VMT	Validation management team
WLR	Within-laboratory reproducibility

About This Report

Authors

Emily N. Reinke[1], Judy A. Strickland[1], Jim Truax[1], David G. Allen[1], Dori Germolec[2], Nicole C. Kleinstreuer[3]

[1]Inotiv, Research Triangle Park, North Carolina, USA

[2]Systems Toxicology Branch, Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

[3]National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Inotiv, Research Triangle Park, North Carolina, USA

Managed project, developed study design, evaluated and interpreted results, and prepared report

Judy A. Strickland, Ph.D., D.A.B.T., Project Coordinator, Validation Management Team Chair

David G. Allen, Ph.D., Principal Investigator

Emily N. Reinke, Ph.D., D.A.B.T.

Jim Truax, M.A.

National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Provided project oversight, evaluated and interpreted results, and reviewed report

Nicole C. Kleinstreuer, Ph.D.

Dori Germolec, Ph.D.

Contributors- Validation Management Team

Simona Bancos, Ph.D., U.S. Food and Drug Administration (FDA), Center for Devices and Radiological Health (CDRH), White Oak, MD, USA

Rakhi Dalal-Panguluri, Ph.D., FDA, CDRH, White Oak, MD, USA

Jill Merrill, Ph.D., FDA, Center for Drug Evaluation and Research, Silver Spring, MD, USA

David Lehmann, Ph.D., U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Research Triangle Park, NC, USA

Inotiv, Research Triangle Park, North Carolina, USA

Edited, formatted, and proofread report

Catherine Sprankle, M.S.

Elizabeth Farley-Dawson, Ph.D.

Christine Burillo-Kirch, Ph.D.

Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Provided contract oversight

Dori Germolec, Ph.D.

Kelly Shipkowski, Ph.D.

Matt Stout, Ph.D.

Pei-Li Yao, Ph.D.

ICF, Reston, Virginia, USA

Provided contract oversight

David F. Burch, M.E.M., Principal Investigator

Jessica A. Wignall, M.S.P.H.

Supported external peer review

Lindsey M. Green, M.P.H.

Aishwarya Javali, M.S.

Peer Review

The National Institute of Environmental Health Sciences Division of Translational Toxicology (NIEHS/DTT) conducted a peer review of the draft "Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers" by letter in August 2024 by the experts listed below. Reviewer selection and document review followed established NIEHS/DTT practices. A summary of the peer reviewer comments has been included as an appendix (Appendix H). The reviewers were charged to:

1. Review the draft "Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers."
2. Comment on whether the draft document is clearly written and objectively presented.

NIEHS/DTT and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods carefully considered reviewer comments in finalizing this report.

Cristina Avonto, Ph.D.
Analytical Chemistry Lead
Bactobio Ltd.
London, UK

Brunhilde Blomeke, Ph.D.
Professor and Chair, Department of Environmental Toxicology
University Trier
Saarland, Germany

Craig Elmets, M.D.
Professor, Department of Dermatology
University of Alabama at Birmingham
Birmingham, Alabama, USA

Kymberly Gowdy, Ph.D., M.S.
Associate Professor and Associate Director, Medical Scientist Training Program
The Ohio State University
Columbus, Ohio, USA

Mahesh Raj Nepal, Ph.D.
Research Scientist
BioCryst Pharmaceuticals
Birmingham, Alabama, USA

Alessandro Venosa, Ph.D.
Assistant Professor, Department of Pharmacology and Toxicology
University of Utah
Salt Lake City, Utah, USA

Publication Details

Publisher: National Institute of Environmental Health Sciences

Publishing Location: Research Triangle Park, NC

DOI: https://doi.org/10.22427/NICEATM-3

Report Series: NICEATM Report

Report Series Number: 03

Official citation: Reinke EN, Strickland JA, Truax J, Allen DG, Germolec D, Kleinstreuer NC. 2025. Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers. Research Triangle Park, NC: National Institute of Environmental Health Sciences. NICEATM Report 03.

Acknowledgments

This project was funded with federal funds from the National Institute of Environmental Health Sciences, National Institutes of Health, under contracts HHSN273201500010C to Integrated Laboratory Systems LLC, a subsidiary of Inotiv, HHSN273201400017C to Burleson Research Technologies, and GS00Q14OADU417 (Order No. HHSN273201600015U) to ICF Incorporated, LLC.

Abstract

This report summarizes the key findings of a validation study for the electrophilic allergen screening assay (EASA). Four laboratories collaborated to conduct a validation study following previously published performance standards (PS) for Key Event (KE) 1 test methods based on the skin sensitization adverse outcome pathway.

The EASA is intended to be a partial replacement for the local lymph node assay (LLNA) in assessing skin sensitization hazard potential. Although multiple Organisation for Economic Cooperation and Development (OECD) test guidelines for skin sensitization have been developed for this purpose, new methods can improve upon testing accessibility, applicability, and throughput. The EASA addresses hapten formation, the molecular initiating event in skin sensitization (also known as KE1). This method has higher throughput and requires less specialized equipment than the existing KE1 assays, the direct peptide reactivity assay and the amino acid derivative reactivity assay (ADRA). All these assays can be used in a weight-of-evidence approach such as an integrated approach to testing and assessment or a defined approach (DA) to assess skin sensitization hazard or potency. An integrated approach to testing and assessment does not have a defined approach for data interpretation and instead relies on weight-of-evidence and expert judgment for decision making. A DA depends on a well-defined, fixed data interpretation procedure applied to data that are generated using a defined set of information sources to derive a prediction. Expert judgment is not needed when using a DA. OECD has accepted and developed Guideline 497, a guideline for DAs for skin sensitization, which incorporates the direct peptide reactivity assay in three different DAs—one for hazard and two for categorization under the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (OECD, 2025b). The ADRA has not yet been incorporated into Guideline 497 and is under evaluation for inclusion at the time of this report. The proposed intent of using DAs for assessing skin sensitization hazard and potency is to replace animal models such as the guinea pig maximization test and the LLNA.

This study assessed transferability of the EASA to naïve laboratories, within- and between-laboratory reproducibility performance of the assays as compared to LLNA reference data, and the applicability domain. PS were developed for methods proposed for inclusion in Test Guideline 442C, which addresses KE1 (OECD, 2019), with an update in 2022, which resulted in a change in the reference chemical list. The changes in the list were made due to changes in the ADRA test guideline, which replaced proficiency chemicals due to addition of the fluorescence detection method and chemical specific variability, affected ADRA outcome. Three proficiency chemicals, which were also in the PS, were changed in both the proficiency chemical list and the PS list. A fourth chemical was replaced because the test concentration in the ADRA changed from 1 mM to 4 mM. The reference chemicals for 2019 and 2022 PS are listed in **Table 1**. This EASA validation study used the 2019 reference chemical list because it was completed prior to the 2022 update.

Table 1. Comparison of reference chemical lists from 2019 and 2022 performance standards

2019 Reference Chemical List	In vivo prediction (2019)	2022 Reference Chemical List	In vivo prediction (2022)
Lauryl gallate	Sensitizer (strong)	Diphenylcyclopropenone	Sensitizer (extreme)
Chloramine T trihydrate	Sensitizer (strong)	Lauryl gallate	Sensitizer (strong)
Metol (4-methylamino phenol)	Sensitizer (strong)	2-methyl-2H-isothiazol-3-one	Sensitizer (strong)
2-Mercaptobenzothiazole	Sensitizer (moderate)	Metol (4-methylamino phenol)	Sensitizer (strong)
Benzyl salicylate	Sensitizer (moderate)	2-Mercaptobenzothiazole	Sensitizer (moderate)
Cinnamaldehyde	Sensitizer (moderate)	Benzyl salicylate	Sensitizer (moderate)
Imidazolidynl urea	Sensitizer (weak)	Imidazolidynl urea	Sensitizer (weak)
Ethyl acrylate	Sensitizer (weak)	Ethyl acrylate	Sensitizer (weak)
Salicylic acid	Nonsensitizer	Salicylic acid	Nonsensitizer
Benzyl alcohol	Nonsensitizer	Propyl paraben	Nonsensitizer
Glycerol	Nonsensitizer	Glycerol	Nonsensitizer
Isopropanol	Nonsensitizer	Isopropanol	Nonsensitizer
p-Benzoquinone	Sensitizer (extreme)	p-Benzoquinone	Sensitizer (extreme)
Dihydroeugenol	Sensitizer (extreme)	m-Aminophenol	Sensitizer (extreme)
Palmitoyl chloride	Sensitizer (extreme)	Palmitoyl chloride	Sensitizer (extreme)
Farnesal	Sensitizer (weak)	Farnesal	Sensitizer (weak)
Benzyl cinnamate	Sensitizer (weak)	Benzyl cinnamate	Sensitizer (weak)
Dimethyl isophthalate	Nonsensitizer	Dimethyl isophthalate	Nonsensitizer
Methyl salicylate	Nonsensitizer	Methyl salicylate	Nonsensitizer
4-Aminobenzoic acid	Nonsensitizer	4-Aminobenzoic acid	Nonsensitizer

Highlighted chemicals indicate the differences between the two lists.

Each participating laboratory completed 10 positive and negative control test plates for each of the EASA probes to determine transferability of the method. Once transfer to participating laboratories was completed successfully, the 20 reference chemicals (**Table 1**) were distributed (blinded) to each testing facility. Twelve of the reference chemicals, which were tested in three qualified tests for each chemical, were used to assess within-laboratory reproducibility. The remaining eight reference chemicals were used to assess between-laboratory reproducibility, along with a two-out-of-three (2o3) concordance determination from each laboratory for the initial 12 chemicals. Both the within- and between-laboratory reproducibility were acceptable to the EASA Validation Management Team, with cumulative scores of 96% and 85% respectively. These scores exceeded the PS requirements of 80%. Performance of the EASA against reference LLNA data was also calculated, with an overall sensitivity of 87%, specificity of 76%, and accuracy of 83%. Although specificity was below the PS criterion of 80%, the EASA predicted sensitizers very well and may overpredict nonsensitizers slightly. However, there were a limited number of nonsensitizer test chemicals and there may have been some applicability domain conflicts (e.g., auto-fluorescence) within the list that limited further the number of chemicals that could be successfully tested. The EASA Validation Management Team found acceptable this

justification for the EASA not meeting the specificity criterion of 80%. Thus, the EASA Validation Management Team found the EASA's reproducibility, sensitivity, and accuracy acceptable for discriminating between sensitizers and nonsensitizers when combined with other information sources, such as those used in DAs or integrated approaches to testing and assessment.

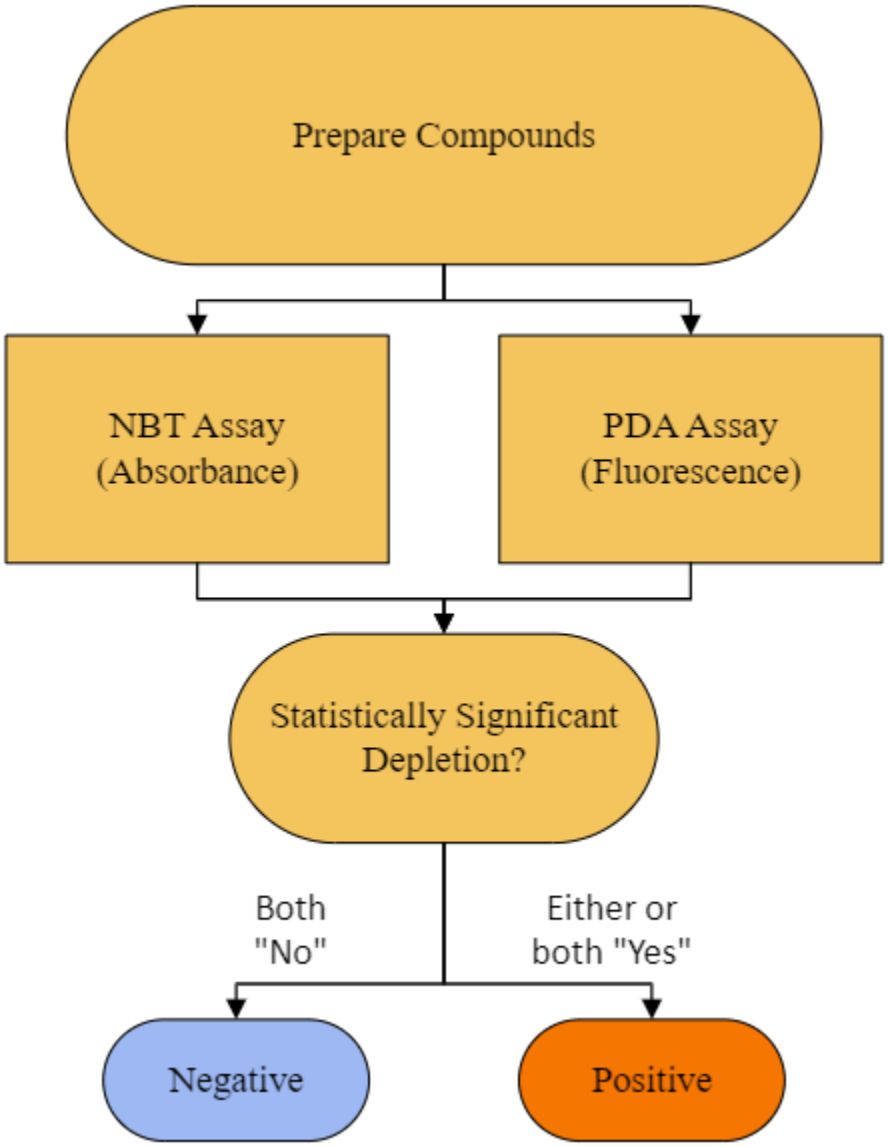
1. Introduction

Every day, dermal exposures to a wide range of chemicals occur in the human population. In a significant portion of this population, exposure can lead to the development of allergic contact dermatitis (ACD), a long-term, potentially debilitating disease that results from a skin-sensitizing event (Alinaghi et al., 2019; Olusegun & Martincigh, 2021). Development of ACD is a major concern for the cosmetics and agrochemical and industrial chemical sectors. For that reason, the potential for a product or ingredient to be a skin sensitizer must be determined through testing prior to availability on the open market or regulatory approval (Daniel et al., 2018). Historically, skin sensitization has been assessed by either a guinea pig maximization test or the local lymph node assay (LLNA), both of which are typically accepted by regulatory agencies worldwide for identifying skin sensitization hazards and represented by test guidelines published by the Organisation for Economic Cooperation and Development (OECD). As we have developed a better understanding of the human immune response and the mechanism by which skin sensitization occurs has been elucidated, we recognize that animal test methods such as the guinea pig maximization test and LLNA do not adequately predict skin sensitization potential (OECD, 2012; OECD, 2021). Additionally, changing ethical mores and legal requirements to limit the use of animals for toxicity testing, particularly for skin sensitization, necessitated the development of non-animal test methods that assess skin sensitization hazard and potency. While multiple test methods have been validated by the OECD over the last several years, including defined approaches (DAs) that combine in chemico*,* in vitro*,* and in silico test methods to classify substances for skin sensitization hazard and potency, there is still a need to expand upon available tests that increase throughput, are more accessible, and expand the applicability domain while decreasing cost as compared to the animal methods (OECD, 2025a, 2025b, 2025c, 2025d).

The primary basis of the development of the various skin sensitization assays has been the adverse outcome pathway (AOP) for skin sensitization (OECD, 2012), a framework to describe the biological key events/steps in the initiation, induction, and elicitation of skin sensitization. There are four key events (KEs) that induce a skin-sensitizing reaction: 1) covalent binding of a small molecule to proteins found in the skin, typically at a cysteine or lysine, which is the molecular initiating event (MIE) and is referred to as hapten formation; 2) induction of inflammatory responses and cytoprotective pathways such as antioxidant response in keratinocytes; 3) mobilization of dendritic cell populations and migration to the lymph nodes, marked via the translocation of specific cellular membrane proteins to the cell surface; 4) T-cell activation and proliferation which results in the adverse outcome of ACD. Existing OECD test guidelines (TG) are in place to assess each of these KEs. TG442C addresses KE1 and currently contains the direct peptide reactivity assay (DPRA), amino acid derivative reactivity assay (ADRA), and the kinetic DPRA (kDPRA) (OECD, 2022c). KE2 is covered by TG442D, containing the KeratinoSens™ and the LuSens assays (OECD, 2025a), and TG442E contains the human cell line activation test, U-SENS™, the IL-8 Luc assay, and the GARDskin, and addresses KE3 (OECD, 2025d). TG 429, TG442A, and TG442B cover KE4 (OECD, 2010, 2018). The test method described herein, the electrophilic allergen screening assay (EASA), is designed to address KE1, covalent binding of a small molecule to proteins found in the skin, the MIE that can lead to ACD if downstream KE steps also occur.

The EASA was originally developed by Chipinda et al. (Chipinda et al., 2010, 2014) at the National Institute of Occupational Safety and Health as a cuvette-based assay and subsequently converted by Petersen, et al. (2022a) to a plate-based assay to increase throughput and accessibility. The U.S. Consumer Product Safety Commission (CPSC) and National Institute of Standards and Technology (NIST) initially developed three assays for use on a 96-well plate, the 4-nitrobenzenthiol (NBT) absorbance assay, and the pyridoxylamine (PDA) absorbance and fluorescence assays (Petersen, et al., 2022a, b). The EASA can rapidly generate skin sensitization data on chemicals in response to public health safety concerns and is accessible to many laboratories, as the detection method requires a plate reader, an instrument commonly found in laboratories. The EASA measures the extinction of probe signal following interaction of a test chemical with either the NBT or PDA probes, with a colorimetric change due to covalent binding of a test chemical to either NBT or PDA or a loss in fluorescence signal following covalent binding to PDA. A decision on skin sensitization hazard is determined based on the test outcome (**Figure 1**), where significant depletion in any probe indicates a skin sensitizer. During this validation study, it was found that the PDA absorbance assay did not improve upon the predictive capacity of the EASA, which resulted in its removal from the approach, leaving the NBT absorbance and PDA fluorescence assays for making hazard classifications.

Figure 1. EASA workflow and decision criteria for identification of electrophilic contact allergen hazard



Abbreviations: NBT = 4-nitrobenzenthiol; PDA = pyridoxylamine.

2. Management of the Study

2.1. Study Objectives

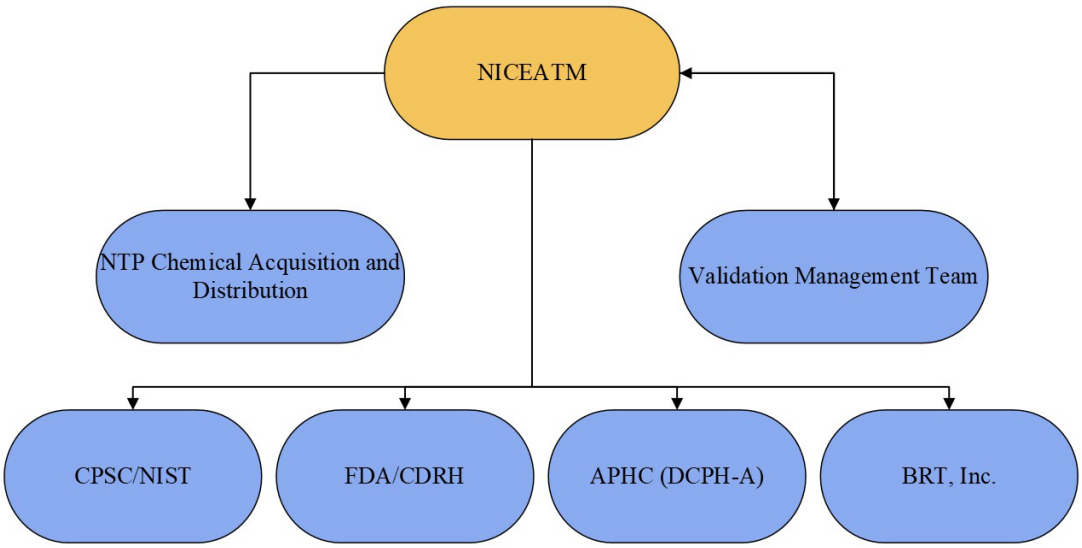
The EASA validation study was launched in 2019 with the following objectives: to optimize the test method protocol from the initial test battery approach for use in any facility, assess intra- and inter-laboratory reproducibility, and assess accuracy for the classification of ACD hazard. This

validation study was run to support an assessment of whether the EASA performs adequately to be included in TG 442C as well as to be evaluated for inclusion in GL 497, Defined Approaches on Skin Sensitization (OECD, 2025b).

2.2. Study Plan

The EASA validation study was based on the "Performance Standards for the assessment of proposed similar or modified in vitro skin sensitization DPRA and ADRA test methods" as described in TG 442C, Series on Testing and Assessment No. 303 by the OECD (OECD, 2019), hereafter referred to as the TG 442C Performance Standards (PS). As the EASA is intended to be incorporated as a KE1 assay under TG 442C, testing against the TG 442C PS was an applicable and efficient approach. The objectives of the study were to optimize the test method protocol, assess inter- and intra-laboratory reproducibility, and to assess accuracy for the classification of hazard, utilizing the 20 reference chemicals from the PS (see **Table 1**). The EASA validation study was coordinated by National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), consulting with the validation management team (VMT) regarding study design, statistical analysis, and protocol alterations. **Figure 2** illustrates the makeup of the reporting structure for the validation study.

Figure 2. Reporting structure for the validation study



Abbreviations: APHC = U.S. Army Public Health Center; BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DCPH-A = Defense Center for Public Health -Aberdeen; FDA/CDRH = U.S. Food and Drug Administration Center for Devices and Radiological Health; NICEATM = National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods; NTP = National Toxicology Program.

2.2.1. Validation management team

The study was coordinated by NICEATM, with Dr. Judy Strickland (Inotiv, contractor supporting NICEATM) serving as the chair and project coordinator. Scientists from NICEATM,

the U.S. Food and Drug Administration (FDA), and the U.S. Environmental Protection Agency (EPA) served on the VMT. See **Table 2** for participating scientists and their affiliations. After testing was completed, Dr. Emily Reinke left the U.S. Army Public Health Center laboratory to join Inotiv and subsequently joined the VMT.

Table 2. Makeup of validation management team

Scientist	Affiliation	Role
Dr. Judy Strickland	Inotiv, contractor supporting NICEATM	Chair/Project Coordinator (retired in 2023)
Mr. Jim Truax	Inotiv, contractor supporting NICEATM	Assistant Project Coordinator (retired in 2023)
Dr. Nicole Kleinstreuer	NICEATM	Director, NICEATM, member
Dr. Dave Allen	Inotiv, contractor supporting NICEATM	Principal Investigator, NICEATM contract, member
Dr. Simona Bancos	FDA/CDRH	Member
Dr. Rakhi Dalal-Panguluri	FDA/CDRH	Member
Dr. Jill Merrill	FDA/CDER	Member (retired in 2022)
Dr. David Lehmann	EPA	Member
Dr. Emily Reinke	Inotiv, contractor supporting NICEATM	Principal Predictive Toxicologist, member (joined June 2022)

Abbreviations: CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; EPA = U.S. Environmental Protection Agency; FDA/CDER = U.S. Food and Drug Administration Center for Drug Evaluation and Research; FDA/CDRH = U.S. Food and Drug Administration Center for Devices and Radiological Health; NICEATM = National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods.

2.2.2. Participating laboratories

Participating federal laboratories were recruited from member agencies of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) interested in the project. CPSC/NIST developed the assay as a follow-on effort from the single-cuvette assay project that preceded this validation study. The EASA validation project testing laboratories met 27 September 2018, with a follow-up kick-off in-person training meeting 8 March 2019. The participating laboratories met a total of 20 times through October 2022. Burleson Research Technologies, Inc. (BRT), a contract laboratory for the National Institute of Environmental Health Sciences, also joined the validation project during this time. Participating laboratories for the validation study and study directors are listed in **Table 3**. Each facility designated one person to conduct the transfer studies and validation studies for consistency. All testing was manually conducted, without the use of a robot.

Table 3. Participating laboratories

Laboratory	Location	Study Director(s)
CPSC/NIST (Lead Lab)	Gaithersburg, MD	Dr. John Gordon Dr. Elijah Petersen
FDA/CDRH	White Oak Campus Silver Spring, MD	Dr. Diego Rua
APHC (DCPH-A)	Aberdeen Proving Ground, MD	Dr. Emily Reinke, replaced by Dr. Valerie Adams (April 2022)
BRT	Morrisville, NC	Dr. Victor J. Johnson

Abbreviations: APHC = U.S. Army Public Health Center; BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DCPH-A = Defense Center for Public Health -Aberdeen; FDA/CDRH = U.S. Food and Drug Administration Center for Devices and Radiological Health.

2.3. Study Design

The EASA validation consisted of two distinct phases, the pre-validation and validation phase. In the pre-validation phase, participating laboratories participated in an on-site training with CPSC/NIST test developers at the NIST laboratories. Following the in-person training, each participating lab conducted positive control (PC) and negative (solvent) control (NC) tests, totaling 10 plates each of NBT absorbance and PDA fluorescence and absorbance assays. The intent of these plates was for each lab to have a reference database of PC and NC data for the development of test acceptance criteria. This also served as the technical transfer portion of the study. The validation phase utilized the reference chemicals set forth in the TG 442C PS (OECD, 2019). Twelve of the specified chemicals were to be tested with three qualified tests (QTs) to assess within- and between-laboratory reproducibility, with the remaining eight chemicals tested once for a total of 20 chemicals to assess between-laboratory reproducibility. One qualified test is comprised of at least two qualified independent repetitions (runs) where the data are concordant. Where the data are not concordant, an additional third independent test must be completed to determine the outcome of the qualified test. The TG 442C PS also specify a minimum of three replicate wells each for the test chemical (TC) and for NCs and PCs. Targeted performance criteria for the reference chemicals were at least 80% for both the between- and within-laboratory reproducibility. Accuracy, sensitivity, and specificity for hazard classification should also be at least 80% with clear rationale for any under-predictions of strong or extreme sensitizers.

2.4. Test Chemical Management

TCs were acquired by the NTP Chemical Repository at NICEATM's request. A chemical safety officer at each facility was identified to receive and complete information on each of the TCs (i.e., physical state, weight/volume, specific density for liquids, and storage instructions). All chemicals were coded at the Chemical Repository and shipped to chemical safety officers at each testing facility. All TCs were weighed into glass vials, capped, sealed with parafilm, and placed in secondary containment plastic bags. Test chemicals were grouped by storage conditions (20°C, +4°C, room temperature) into metal canisters. Safety officers were instructed on

cataloging, storage, and handling for each chemical. Sealed safety data sheets were provided for all 20 TCs, with a separate code breaker spreadsheet emailed to safety officers. Chemical disposal was under the purview of the safety officer at each facility, following local regulations for appropriate disposal. Chemical coding information for each lab is provided in **Appendix A**.

2.5. Data Management

The lead laboratory, CPSC/NIST, designed check sheets for use by participating labs for preparation of reagents, study conduct, and data analysis. Data was entered into Microsoft Excel™ data calculator files, prepared and locked by CPSC/NIST. A worksheet for quality control (QC) chart data was provided by the data calculator for entry into a Microsoft Excel QC chart file. QC values were copied and pasted between the two files as specified in the worksheet calculations to determine if an assay met acceptance criteria and to make determinations on TC outcomes. All electronic worksheets were stored on servers as specified by each individual participating laboratory, and electronic copies were provided to NICEATM for final collation and evaluation. Paper check sheets were stored as specified by each individual testing facility.

2.6. Validation Study Statistical Analysis of the Data

2.6.1. Data analysis

Acceptance criteria (see Section **4.4**) were established to determine if a run was successful and could be included in statistical analyses, although all records were available to allow for analysis of failed run occurrence and frequency. Only acceptable runs that met acceptance criteria were included as part of the evaluation of the validation study. Overall, for the conduct of this assay the failure rate ranged from 0 to 18% depending on probe and facility, with an average overall failure of 4% for all plates evaluated. These analyses included assessment of the inter- and intra-laboratory reproducibility (for reliability), dependent on concordance of the data with the acceptance criteria set out in the TG 442C Performance Standards (OECD, 2019).

3. Test Method Rationale/Description/Definition

3.1. Intended Purpose of the Test Method

The EASA is intended to be applied as part of a weight-of-evidence approach for the determination of skin sensitization hazard; this can be part of an integrated approach to testing and assessment, relying on expert judgment to make a final decision, or as part of a DA, such as those found in TG 497. The intent behind the development and validation of this method is to provide a mechanism to assess the potential for skin sensitization in a more expedient, economical manner using common laboratory equipment that also has the benefit of reducing or eliminating the need for animals in the prediction of skin sensitization, which can lead to ACD, a potentially long-term, debilitating condition. The method is designed to address the MIE for the skin sensitization AOP, e.g. the binding of the skin-sensitizing chemical to a protein.

3.2. Evidence Demonstrating Need of the Test Method

Multiple regulatory requirements have been put in place in recent years that encourage the use of new approach methodologies (e.g., Frank R. Lautenberg Chemical Safety for the 21st Century Act for new chemicals or EPA interim science policy for pesticide products) or ban the use of animal methods for specific consumer products (e.g., cosmetics products in the European Union, 2009/1223/EU, or California, SB 1249) (EPA, 2016, 2018; European Union, 2006, 2009; Nel & Malloy, 2017). Successful validation of the EASA could address regulatory requirements both in the U.S. and internationally by providing a non-animal method to help predict skin sensitization hazard within a DA or integrated approach to testing and assessment. Additionally, building upon the understanding of the mechanisms of skin sensitization, the EASA would provide a higher throughput, more efficient method to detect the MIE in the skin sensitization AOP. Current limitations on solubility for other similar methods may also be overcome by the EASA due to its ability to be conducted with a range of solvents.

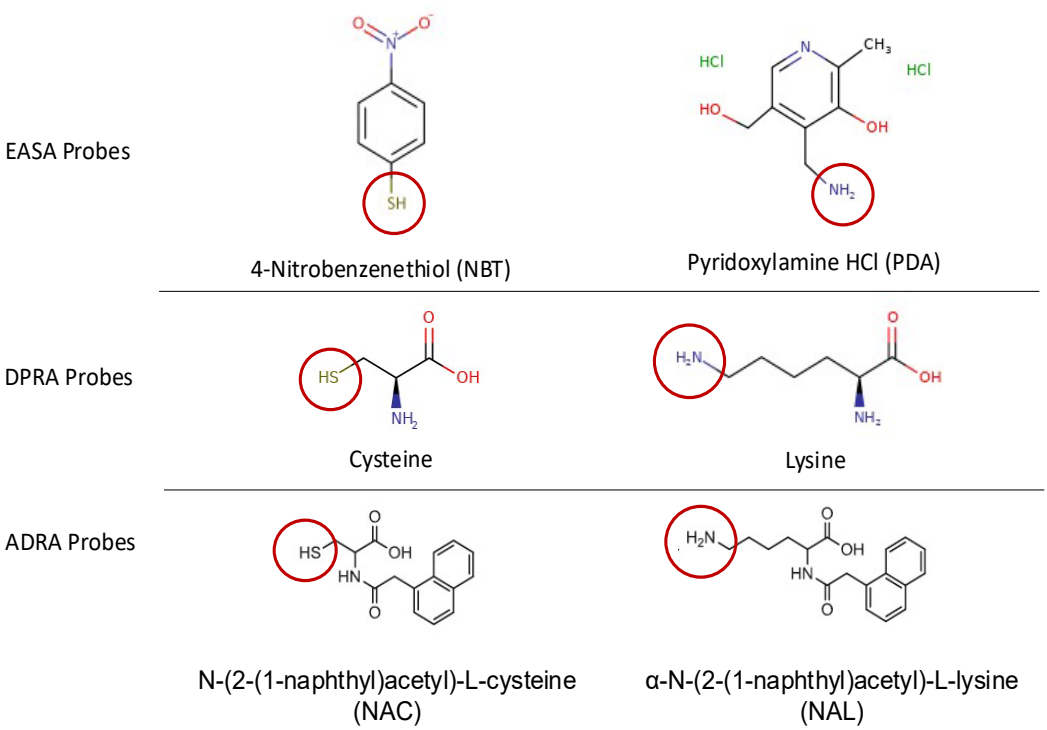
While each of these assays has noted strengths and good predictive capacity, they also have weaknesses that limit their applicability. For the KE1 assays DPRA, ADRA, and kDPRA, throughput is a primary limiting factor (OECD, 2025c). They are time-intensive, with a limited number of chemicals that can be tested in any single run, and each requires significant machine up-time (maximum of 30 hours) for each run, with multiple runs required for a complete test. With the incubation period, each test can take up to 54 hours from setup to final sample (Petersen, et al., 2022a). Additionally, all three of these assays require specialized equipment and training that may not be found in all research or testing facilities. The EASA, however, has higher throughput, with the potential to assess up to seven chemicals in a single run (versus two in DPRA/ADRA and one in the kDPRA), requires less than an hour of instrument time per run, and can be completed on a standard platform plate reader, which can result in an overall cost savings. This makes the EASA more accessible and with higher throughput than the existing KE1 assays. As with the DPRA, ADRA, and kDPRA, the EASA also does not have metabolic capacity to detect pro-haptens. This is a major limitation of not only the in chemico KE1 assays but is a consideration of the cell line-based assays as well (OECD 2022a, 2022b, 2022c).

3.3. Biological and Mechanistic Relevance of the EASA

The principle behind the EASA is the same as for the other validated KE1 assays: an electrophilic small molecule (sensitizing chemical) interacts with a nucleophile (typically a cysteine or lysine on a protein). However, the EASA is designed such that that the relative binding of a potential chemical allergen is based on the hard/soft (Lewis) acid/base concept (Lopachin et al., 2012). This concept says that soft electrophiles form covalent bonds more quickly and strongly with soft nucleophiles, while hard electrophiles similarly bind better to hard nucleophiles. **Figure 3** shows the structure of the EASA probes in comparison to DPRA and ADRA nucleophiles. Here the NBT is used as the cysteine surrogate, or soft nucleophile, and PDA is the lysine surrogate, or hard electrophile. Covalent binding of an electrophilic allergen to the amine or thiol on these probes (red circles) produces a shift in the absorbance and/or fluorescence. The shift (loss) of absorbance/fluorescence is directly related to the allergen's chemical reactivity and can be directly monitored continuously or through endpoint measures.

Additional benefits of utilizing chemical probes include lower concentrations of TC to decrease the likelihood of insolubility and precipitation, shorter assay times, and lower cost.

Figure 3. KE1-based nucleophiles



Abbreviations: ADRA = amino acid derivative reactivity assay; DPRA = direct peptide reactivity assay; EASA = electrophilic allergen screening assay.

3.4. Development of the Test Method

As described above, the hard/soft acid/base principle is the basis for the EASA test method. The method was originally developed as a cuvette-based method, where there was a positive correlation between the EASA and LLNA EC3 values (TC concentration giving a stimulation index of 3), and neither probe reacted to nonsensitizers (Chipinda et al., 2014). In 2012, the NBT and PDA assays were nominated to ICCVAM for a validation study as the EASA using the cuvette-based approach. This approach limited throughput and accessibility, due to the need for cuvette-based spectro/fluorometers and a limited number of cuvettes that could be loaded into any instrument at a given time. However, the assay performed well in an initial validation study, with high intra- and inter-laboratory concordance (internal data). With the promising results from the cuvette-based Phase 1 trial, CPSC/NIST collaborated to develop a plate-based EASA, in order to increase the controls, throughput, and efficiency of the assay. Using a measurement science process, robustness testing was conducted to minimize sources of variability that affected assay performance (Petersen et al., 2022a). Cause-and-effect analyses were also conducted throughout the design process. Sources of variability assessed in the development of the assay are highlighted in **Figure 4** and covered in greater detail in the peer-reviewed publication about

the development of the 96-well EASA (Petersen et al., 2022a). Assessment of different plate types was conducted to determine which types of plates would have resistance to acetonitrile (ACN) and the appropriate transparency to detect fluorescence signals at specified wavelengths. Quartz plates were considered but based upon the need for standardized washing protocols and the high cost of a single plate and lid, other options were investigated first. Flat-bottomed cyclic olefin copolymer 96-well plates were found to have adequate resistance to degradation by ACN (no detectable change in absorbance or fluorescence signal over a 120 min period) and appropriate transparency at a reasonable cost-rate per plate. Different sealing tapes were also assessed for resistance to ACN vapors until a suitable option was found. Consumable details are provided in the protocol (**Appendix B**).

Figure 4. Potential sources of variability assessed in the development of the 96-well plate EASA

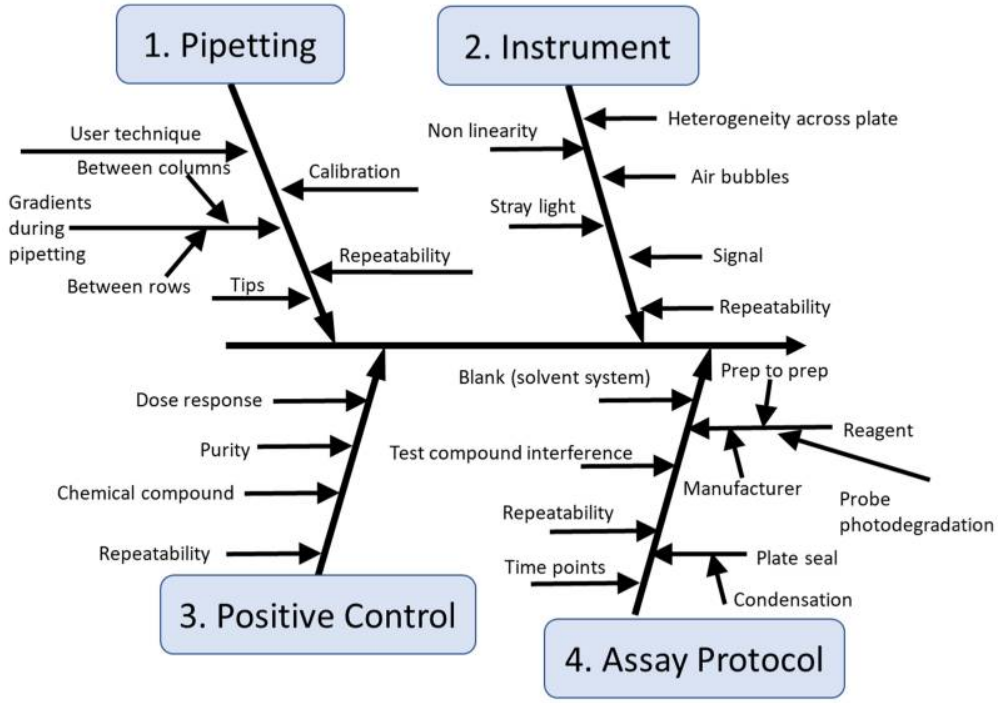


Image from: Petersen et al., 2022a.

In addition to plate type selection, CPSC/NIST assessed multiple different factors to increase robustness of the assay. Primary areas of concentration were in optimizing the exposure duration, PC selection and concentration, stability of probe molecules, plate reader homogeneity and impacts of pipetting direction, potential for TC interference, bias from bubble formation in wells, and expanding solvent types to include polar and semi-polar solvents. These assessments revealed biases within the cuvette-based assay that could be mitigated on the 96-well plate. The NBT was found to rapidly degrade under white light, an issue that was remedied using red light when working with this probe. Condensation on the sealing tape and effects of plate cooling or heating were evaluated, as were the differences in utilizing polystyrene or polypropylene reservoirs for chemical stocks when plating. Full details of these assessments can be found in the EASA development publication (Petersen et al., 2022a).

The results of these assessments led to a plate design that contained several in-process control measurements, such as measurements for within and between columns pipetting variability, enough wells to provide statistical robustness and subtract for background (solvent system [SS; 1:1 ACN:phosphate buffer] only) and wells to assess for TC interference (test chemical without probe) (**Figure 5**, **Table 4**).

A dose response for the PC was also added to determine sensitivity of the assay for each run. Thus, the final plate layout includes an adequate number of blanks (SS only), NC wells (probe only), PC replicates, and TC replicates with and without probe. The plate was designed for up to seven TCs to be assessed in both an NBT and a PDA assay on the same day with multiple sets of NBT and PDA runs possible in the same day. A feasible approach allows up to 14 chemicals to be tested per day. To control for physical interference, such as bubbles in the wells, an additional absorbance measurement at 680 nm, outside the absorbance spectrum of the probe molecules, was added during the initial 5-minute incubation. Finally, assay duration was determined as a combination of the following factors: onset of tape fogging, probe signal depletion in the NC wells, and the rate of PC percent depletion. These factors were evaluated separately to determine when probe signal began to decrease, typically after 50 minutes, and when tape fogging or condensation occurred, which was typically after 60 minutes. With these factors in mind, it was determined that the 50-minute time period was appropriate (Petersen et al., 2022a). Acceptability of a test run relied on each of these in-process controls meeting specific statistical criteria, which are calculated in the data calculator file. The in-process controls include the coefficient of variation (COV) of the NC, and the half-maximal inhibitory concentration (IC_{50}) value of the PC must not fall outside three times the standard deviation for the mean for each of these, based on pre-validation plate data. Over time, this has been amended to be 1.5 times the interquartile range (IQR) and ultimately amended to three times the IQR. The means of the SS (NC/PC blank), NC, and individual PC concentrations for each run are also assessed and flagged with a “Caution” if they fall outside three times IQR of the pre-validation data. This “Caution” indicates that there may be an issue with buffers or the probes and should be monitored.

Prior to a test facility conducting TC assessment, the QC parameters must be established. Ten NC/PC plates for both NBT and PDA are tested to provide the QC comparisons for each TC and plate for run acceptability. During this stage, PDA concentration is also titrated to insure it does not cause overflow reads on the plate reader. This was the pre-validation step in the validation study.

Figure 5. 96-well plate design to cover in-process controls and positive control dose response

- NC/PC Blank wells (40 µl ACN + 160 µl SS without Probe)

- PC (40 µl PC in ACN + 160 µl SS with Probe)

- NC (ACN) (40 µl ACN + 160 µl SS with Probe)

- TC (40 µl TC in ACN + 160 µl SS with Probe)

- TC Blanks (40 µl TC in ACN + 160 µl SS without Probe)

- Not used– no additions

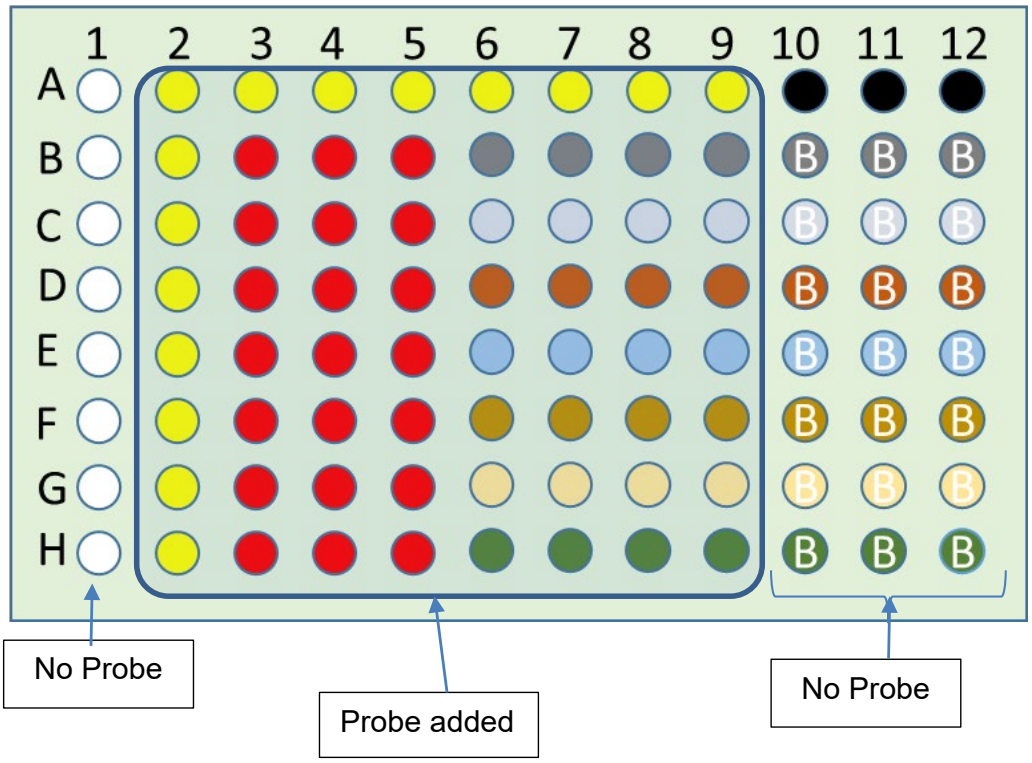


Figure from Petersen et al., 2022a
Abbreviations: ACN = acetonitrile; NC = negative control; PC = positive control; SS = solvent system TC = test chemical.

Table 4. Specifications for the EASA assay based on the robustness testing

	Name	Description	Test wells	Specification	What happens if specification is not met
1	Bubble analysis	Tests for bias from inadvertent bubbles in sample well	All	Absorbance at 680 nm ≥ 0.081 (unless observed in all wells for a TC).	Remove well from analysis.
2	Interference analysis	Tests if a TC causes a change in the absorbance or fluorescence in the absence of the probe molecule	B10-H12	Interference is defined two ways. First, potential interference is detected if the mean TC interference wells (Column 10–12) plus one standard deviation is less than the mean of the NC wells minus one standard deviation. This comparison evaluates if the absorption or fluorescence values decrease due to interference. In some cases, a compound can absorb at the fluorescence test wave lengths, resulting in a lower TC Blank value. Second, an Excel-based one-sided t-test is performed to see if there is a significant difference in the NC wells and the TC interference wells. If the calculated p-value is less than the p-limit, the compound is considered to have interfered with the reading.	The same calculation is performed, but a note is made on the output file that interference was observed. However, it is possible to dilute the compound to minimize interference if desired.
3	Solvent system wells	Evaluate if solvent system wells' mean value varies from historical range	A1-H1	Compare solvent system wells to historical range. Plate is an outlier if the mean solvent system well value exceeds three times the standard deviation value of the historical mean values.	Repeat plate.
4	NC wells	Evaluate if NC well's mean varies from historical range	B2-H2, A3-A9	Compare NC wells to historical range. Plate is an outlier if the mean NC well value exceeds three times the standard deviation value of the historical mean values.	Repeat plate.
5	PC wells	Evaluate if PC IC_{50} value varies from historical range	B3-H5	Compare PC IC_{50} values to historical range. Plate is an outlier if the PC IC_{50} exceeds three times the standard deviation value of the historical IC_{50} values.	Repeat plate.
6	Number of wells	Evaluate if too many wells have been removed for different components of the assay for the result to be reliable	All wells except A10-12	There is a minimum number of wells required after wells are removed (as a result of outlier results, or from an overly strong signal extending beyond the instrument dynamic range).	Repeat plate or exclude results for a specific TC.

Abbreviations: IC_{50} = half-maximal inhibitory concentration; NC = negative control; PC = positive control; TC = test chemical.

During test method development, an initial screen of 92 TCs was completed to evaluate the performance of the assay. Results were statistically analyzed via Bayesian and frequentist

approaches. Endpoints assessed were PDA fluorescence and absorbance and NBT absorbance. PDA absorbance was the least predictive, with the fewest compounds showing a significant amount of probe depletion; inclusion of this endpoint did not enhance performance of the assay beyond the NBT absorbance and PDA fluorescence assays. This prompted the removal of the PDA absorbance assay from the suite. Predictions from the 92-test-chemical set were compared to compiled LLNA data where available, with a best prediction of accuracy of 77% (49/64) using the frequentist approach with an $\alpha = 0.005$. The sensitivity was 60% (9/15) and specificity was 82% (40/49). The DPRA for the same set of chemicals had an accuracy of 77% (34/44), sensitivity of 85% (11/13), and specificity of 74% (23/31). Of the chemicals tested, 32 had human data available in a human predictive patch test database (Strickland et al., 2023). The EASA had a sensitivity of 81%, specificity of 27%, and accuracy of 63% against the human data. The DPRA against the human data set (31 chemicals) had a sensitivity of 86%, specificity of 90%, and accuracy of 87%. For reference, five of the 20 PS chemicals were included in the 92 chemical test set. A full analysis of the chemical performance is available in the publication (Petersen et al., 2022a).

4. Assay Conduct and Development

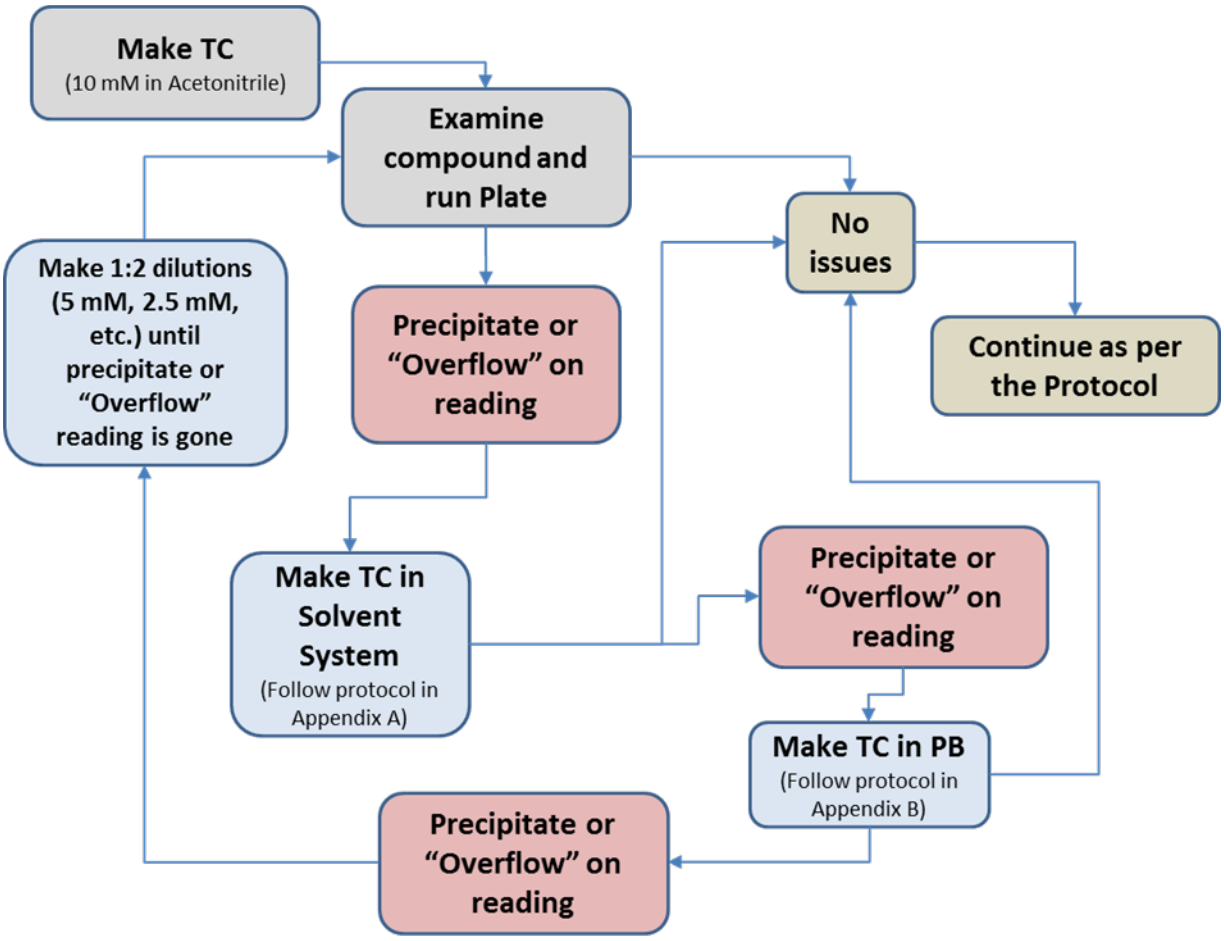
4.1. Protocol and Data Calculator/QC Chart Development

The initial validation study protocol was written following the development of the 96-well assay by CPSC/NIST and can be found in **Appendix B**. Subsequent version changes are also included in **Appendix B** with specifications on what changes were made. In order to conduct the statistical analysis, a set of three Excel worksheets were designed for each of the labs to work from, which included a QC chart and two data calculators, one each for NBT absorbance and PDA fluorescence assays (**Appendices C and D**). The data calculators process the raw absorbance and fluorescent measurements to provide positive or negative test outcomes for skin sensitization potential. The QC charts track the performance of the test acceptance criteria (**Appendix E**). Assays were conducted using robust quality management systems to assure the quality and integrity of the data, and check sheets were provided for preparation of stock reagents such as buffers, solvents, SS, TCs, PCs, and chemical probes (**Appendix F**).

4.2. Testing Reagent Preparations and Testing Flow

Stock reagents such as buffers, solvents, and PC chemicals can be prepared in advance of the conduct of the assay, with storage times up to 30 days at room temperature for SS, PDA stock, and both benzyl bromide and glutaraldehyde PCs. NBT is prepared under red light, aliquoted, and stored at -20°C for up to seven days. On the day of testing, TCs are weighed and prepared as 10 mM stock solutions in the appropriate solvents. This is typically ACN unless there are solubility incompatibilities, at which point, the decision tree is followed (**Figure 6**). Once TCs are prepared, the first plate is set up as per the protocol. During the initial 5-minute incubation, a bubble test is conducted at 680 nm. After the bubble test, absorbance (412 nm for NBT) or fluorescence (324 nm ex/390 nm em for PDA) are measured at 15-minute intervals up to 50 minutes.

Figure 6. Test chemical suitability assessment flow chart



Abbreviations: PB = phosphate buffer; TC = test chemical.

4.3. Data Analysis

A positive call in the assay is not based on a threshold response as in other KE1-based assays but on statistical difference from the NC utilizing a modified t-test. The use of a modified t-test accounts for variability and uncertainty within individual groups of readings, such as the NC, PC, TC, and all blanks. Each of these groups will have different sources of variability and uncertainty in each well. The use of a statistical approach allows for quantification of all sources of uncertainty and variability in making positive/negative determinations and is thus more comprehensive and statistically robust. If the uncertainty from the test compound interference wells was not used, it could lead to a significant underestimate of the total uncertainty and an increased likelihood of false positive results. This statistical approach is consistent with best practices for uncertainty evaluation measurements.

Following a test, the raw data from each read (Bubble, 5-, 20-, 35-, and 50-minute) are entered into the "Raw Data" tab of the NBT or PDA data calculators, along with relevant test information such as date, chemicals tested, testing concentration, solvent, and the identification

of the person conducting the study. The data calculators automatically calculate test acceptance criteria for each of the TCs, IC_{50} for the PC and other QC parameters (**Table 5**). These are compared against the static QC chart data to determine test acceptance. The TCs are also evaluated for interference, adequate well numbers, and other performance criteria and a hazard call is determined based on the outcome of the t-test. The data calculators and the QC chart are necessary to determine if all the QC parameters are met and to make a call for whether a chemical is or is not a sensitizer.

4.4. Acceptance Criteria

Acceptance criteria (**Table 5**) were established to determine if a run was successful and could be included as part of the predictive decision tree as described in the EASA protocol. A run was assessed statistically to determine if it met the acceptance criteria, based upon outcomes by both the data calculators and the locked QC chart worksheets. Only successful runs that fulfilled these criteria were included as part of the evaluation of the validation study.

Table 5. Acceptance criteria for test run

QC Parameter	Minimum Criteria
680 nm "Bubble Test"	Minimum number of wells required to pass: NC/PC Blank: n = 6 Negative Control (NC): n = 12 Positive Control (PC): n = 2 Test Chemical Replicates: n = 3 Test Chemical Blanks (TC Blanks): n = 2
NC Coefficient of Variation	Within 3x IQR*
Forecast IC_{50} of PC standard curve	Within 3x IQR*

Abbreviations: IC_{50} = half-maximal inhibitory concentration; IQR = interquartile range based on the pre-validation data for each laboratory; NC = negative control; PC = positive control; QC = quality control.
*3x IQR is established by the Excel function Quartiles for each QC parameter.

4.4.1. Bubble test

Prior to the absorbance or fluorescence measurements for NBT or PDA, an absorbance measurement at 680 nm is conducted to determine if any physical interference, such as bubbles or precipitate, is present in the well. A single well with an absorbance above 0.085 is an outlier and will be excluded for analysis on the calculation sheet. Colored compounds that absorb light at the lower visible wavelengths, such as red compounds, may have a high absorbance in this test. In the case where all seven test chemical absorbance measurements (TCs and TC blanks) are above 0.085, these will not be excluded from analysis.

For each test grouping (NC, PC, TC, blanks, and TC blanks) there are a minimum number of wells that must pass the bubble test in order for the plate data to be acceptable (**Table 5**). If the NC/PC blank, NC, or PC fail these requirements, the entire plate fails the run; if the TC fails, only that specific chemical fails the run.

4.4.2. Baseline QC parameters

Once an assay run passes the bubble test, the QC charts file calculates whether the NC/PC blank, PC, and NC results are acceptable. QC acceptance parameters for each of the participating laboratories were established during the pre-validation study. The pre-validation included a total of 10 PC/NC plates for both NBT and PDA. QC parameter data (**Table 5**) from each of the pre-validation runs were copied from the NBT or PDA test calculator into the appropriate tab on the QC chart. These data points included absorbance or fluorescence values for each well for NC/PC blank, NC, the NC COV, and the average percent depletion for each concentration of the PC. For each pre-validation plate, a mean and standard deviation were calculated for each endpoint, as well as a forecast IC_{50} for the PC. From these individual values, the overall mean, standard deviation, and quartiles were calculated. The upper and lower limit, defined by the Excel Quartile function, and specified as 3x the IQR for the NC/PC Blank, NC, NC COV, Midpoint, Forecast IC_{50}, and each concentration of the PC are copied into the NBT or PDA test calculators for use with each TC run. For a run to be acceptable the NC COV and forecast IC_{50} must fall within 3x IQR. Additional data flags for the mean NC/PC blank, NC, PC midpoint, and PC dose-response curve are available but are not required for an assay run to pass. They may provide insight into performance of the assay.

5. Problems Encountered During Validation Process

The pre-validation of the EASA was launched in April 2019, with labs completing their 10 pre-validation plates throughout the summer and fall of 2019. During the pre-validation period, additional analyses were conducted on existing 96-well plate data from the assay development studies. Based on these analyses, the VMT dropped the PDA absorbance assay from the validation study. The PDA absorbance assay did not add any robustness to the predictivity of the EASA beyond what the PDA fluorescence and NBT absorbance assays were capturing. Concurrent to this analysis and the pre-validation study, shipments of chemicals were prepared by the NTP Chemical Repository, with labs receiving reference chemicals in February 2020. Testing was held until receipt of a revised protocol, data calculators, and QC charts to accommodate the VMT's decision to change the calculation of the acceptable QC ranges from three times the standard deviation to 1.5x IQR. Additional analyses of the pre-validation plates by NIST revealed that one test facility, Department of Defense (DoD), had a bias toward increasing fluorescence output from left to right on the plate, which required additional investigation. Shortly after these assessments, the COVID-19 pandemic shut down all participating laboratories for non-essential testing/use, therefore causing a delay in the start of the validation phase.

The FDA and DoD laboratories were allowed to open for restricted on-site use in the summer of 2020. FDA began testing coded reference chemicals while DoD began troubleshooting the left to right fluorescence bias for PDA. FDA completed testing of reference chemicals by the end of August 2020. The plate reader calibration was confirmed to be appropriate, and the machine performing within manufacturer specifications. DoD found that the detected PDA signal was very high bordering on being outside the upper range of detection on their plate reader, due to a high sensitivity of the plate reader. A 50% reduction in PDA concentration brought values for PDA fluorescence in line with other test facilities, and well below the upper limit of the plate

reader. This prompted a revision to the protocol to allow for optimization of the PDA concentration/signal. The DoD facility completed a new pre-validation set of 10 plates for the PDA fluorescence assay.

Another delay for reference chemical testing at the DoD and CPSC/NIST laboratories occurred in March 2021 when it was found that some of the TCs had expired. The NTP Chemical Repository purchased fresh chemicals and then shipped new aliquots of coded reference chemicals to CPSC/NIST, DoD, and an additional laboratory, BRT, in June 2021.

While testing proceeded for each of the labs, additional supply chain and plate consistency issues emerged. Due to the pandemic, consumable supplies, primarily plastics, were extremely difficult to source, with plate availability the most impacted. This led to a delay in testing by both BRT and CPSC/NIST. Additionally, the fluorescent background on the clear plates in newer lots increased, causing plates to fail when compared to the historical pre-validation QC parameters. Laboratories screened older lots of plates for fluorescence background similar to the pre-validation plates and set these aside for PDA use only, or black-walled plates were substituted after verification that they were interchangeable. These plates were shared between labs to ensure testing could continue. NBT was also found to have lot-to-lot variability at both CPSC/NIST and DoD, with lots both above and below the QC parameters for the NC wells. Some variation was due to tighter QC criteria (1.5x IQR vs 3x IQR) but was mostly due to variability between the lots. The stability of opened containers of probes was a concern, due to stocks sitting in storage between the end of the pre-validation phase and the start of validation due to the COVID-19 shutdown period. The lead lab determined that in the future, reagents should be assessed for concordance with the QC parameters and if found to be out of specifications, new QC parameters specific to the current lot/time period should be determined. These steps can be taken upon receipt of a new lot, or after a longer gap in testing.

During testing of the 20 reference chemicals, there was an HVAC (heating, ventilation, and air conditioning) issue with the DoD facility that resulted in very high ambient laboratory temperatures. The standard temperature of the facility in summer/fall is typically 19-21°C. During this period, the facility reached upwards of 25°C with the chamber temperature in the plate reader reaching up to 27°C depending on time of day and machine up-time. NBT was more sensitive to temperature fluctuations, with an approximately ± 2°C margin of error based upon the pre-validation QC parameters. The PC dilution series and forecast IC_{50} were the most frequently affected. PDA did not appear to be affected by temperature fluctuations.

The other issue that occurred while testing the 20 reference chemicals was that the FDA laboratory had difficulty meeting the test acceptance criteria established by the pre-validation study. Although the NBT PC IC_{50} was within control limits and the NC COV was consistently tight, the NBT blanks and NCs were consistently just above the cut-off for the 3x IQR. For the PDA assay, several blanks were also outside the 3x IQR, but other NC parameters were largely acceptable; this was likely due to plate lot variability. The PDA PC IC_{50} was outside the 3x IQR established during the pre-validation as the concentration of PDA used during the pre-validation study was twice what it should have been (FDA did not use the preparation check sheet offered by the lead lab) and the forecast IC_{50} range was not corrected for the validation study. When PDA concentrations were appropriately adjusted, the forecast IC_{50} for each of the test plates fell

within the newly established 3x IQR. The issues with passing QC criteria for the NBT and PDA plates were likely not related, but some potential causes for both sets of plate failures are listed in **Table 6**.

Table 6. Potential causes for FDA test failures

NBT Causes	PDA Causes
Reagent expiration	Incorrect PC concentration
Improper reagent concentrations (probe)	Lot-to-lot variation of plates (lot numbers not recorded)
Equipment failure – new multi-channel pipettor may not have been calibrated appropriately	

Abbreviations: NBT = 4-nitrobenzenthiol; PC = positive control; PDA = pyridoxylamine.

6. Transferability

In order to meet the goals of the validation study, transferability of the assay needed to be assessed. Assessment of transferability was accomplished through an in-person training and the completion of 10 pre-validation NC/PC plates for NBT absorbance, PDA fluorescence, and PDA absorbance (later dropped from study). Successful execution of the 10 pre-validation plates would demonstrate that each facility has successfully learned the conduct of the assay. This stage also allowed for optimization of the test protocol and calculations worksheets based on feedback from each of the participating laboratories. The lead laboratory ensured that all test facilities were provided with adequate support during this transfer stage to conduct the assay, and then throughout the validation process via the NICEATM management team. As necessary changes to the protocol or calculations approaches were noted, the lead lab conducted the feasibility studies (changes in statistical assessment) to determine the best approach, and this was then communicated out to participating laboratories, and new calculations sheets were provided.

6.1. Training at NIST

In March 2019, prior to the pre-validation study period, the NICEATM study management team and the laboratory leads from FDA and DoD met with the lead laboratory (CPSC/NIST) at the NIST testing facility. This meeting was to convey the premise of the assay and to have hands-on training in the conduct of the test and utilizing the calculator spreadsheets and QC charts. Items covered in this training were an overview of NIST, the validation study design, the development of EASA, statistical assessment and process control measurements, protocol and worksheets, and an in-person laboratory demonstration of the assay. After BRT joined the validation study, the same training was provided virtually to participants at BRT due to COVID-19 restrictions.

6.2. Pre-validation PC/NC Plates to Establish Baseline QC Parameters, Determine Transferability

Prior to testing the reference chemicals in the EASA, each lab established baseline QC parameters for use to evaluate qualifying test runs. These summary data included upper and lower limits for NC/PC blank, NC, and NC COV. For the PC, the midpoint, forecast IC_{50} and

dilution series upper and lower limits were also established. For a test run to pass QC, only the NC COV and forecast IC_{50} must be within the bounds established by the QC charts; the remaining data provide additional information on the quality of the run and if there are any issues to be concerned with, such as the kinetics of the PC changing due to stability of the chemical or external factors such as temperature. Individual QC charts and all test data for each of the participating laboratories are provided in Appendix G**.**

6.2.1. Individual laboratory performance

6.2.1.1. Lead lab (CPSC/NIST)

The lead laboratory conducted their NBT pre-validation plate runs over three days, establishing NBT QC data by the end of October 2021. The PDA QC data had to be re-established three times throughout the validation process, with the original 15 pre-validation plates completed between March and May 2019. In early February 2022, 10 new control plates were tested, followed by 10 new plates in July 2022. The QC plate criteria for PDA were re-established because test substances were failing test acceptance criteria due to PDA lot-to-lot variability as well as lots of plates that had different background fluorescence of the blank wells. Thus, the lead lab needed to establish new control data so that TC runs would pass test acceptance criteria. As a result, future iterations of the protocol will specify when new criteria should be established for both NBT and PDA, based upon the performance of the probe or PC as well as plate background readings. **Table 7** shows the summary QC Report Data for CPSC/NIST.

Table 7. CPSC/NIST baseline parameters

NBT absorbance QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Benzyl bromide (PC) 0.600 mM	Benzyl bromide (PC) 0.300 mM	Benzyl bromide (PC) 0.150 mM	Benzyl bromide (PC) 0.075 mM	Benzyl bromide (PC) 0.038 mM	Benzyl bromide (PC) 0.019 mM	Benzyl bromide (PC) 0.009 mM
Upper Limit	0.072	0.863	3.48%	54.1%	0.1131	99.3%	97.1%	78.6%	52.5%	29.2%	17.7%	10.0%
Lower Limit	0.067	0.559	-0.94%	50.4%	0.0848	96.1%	87.4%	64.5%	37.4%	21.5%	10.1%	3.3%

PDA fluorescence QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Glutaralde-hyde (PC) 0.0200 mM	Glutaralde-hyde (PC) 0.0100 mM	Glutaralde-hyde (PC) 0.0050 mM	Glutaralde-hyde (PC) 0.0025 mM	Glutaralde-hyde (PC) 0.0013 mM	Glutaralde-hyde (PC) 0.0006 mM	Glutaralde-hyde (PC) 0.0003 mM
Version 1 Upper Limit	5416.3	53487.5	6.78%	57.8%	0.0035	103.9%	100.6%	95.6%	65.2%	36.7%	20.8%	15.1%
Version 1 Lower Limit	4443.8	32206.6	-1.74%	44.7%	0.0022	91.1%	88.7%	69.9%	32.7%	12.4%	1.5%	-5.6%
Version 2 Upper Limit	16107.5	55741.8	4.98%	52.6%	0.0031	99.8%	97.7%	88.4%	51.8%	32.1%	15.8%	9.6%
Version 2 Lower Limit	10065.8	33521.7	-1.45%	50.0%	0.0027	95.2%	91.7%	74.2%	43.9%	17.0%	7.4%	1.0%
Version 3 Upper Limit	13342.6	55377.2	10.91%	61.1%	0.0032	104.0%	100.4%	91.7%	84.7%	51.0%	29.8%	25.8%
Version 3 Lower Limit	9357.9	24926.4	-4.23%	47.5%	0.0016	93.3%	92.6%	86.8%	39.5%	16.7%	4.5%	-5.0%

Abbreviations: COV = coefficient of variation; IC_{50} = half-maximal inhibitory concentration; NBT = 4-nitrobenzenthiol; NC = negative control; PC = positive control; PDA = pyridoxylamine; QC = quality control.

6.2.1.2. FDA

The FDA-CDRH facility completed the pre-validation process throughout November 2019, finishing with 14 NBT plates and 12 PDA plates to establish the baseline QC parameters. Their baseline QC parameters are shown in **Table 8**.

Table 8. FDA baseline QC parameters

NBT absorbance QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Benzyl bromide (PC) 0.600 mM	Benzyl bromide (PC) 0.300 mM	Benzyl bromide (PC) 0.150 mM	Benzyl bromide (PC) 0.075 mM	Benzyl bromide (PC) 0.038 mM	Benzyl bromide (PC) 0.019 mM	Benzyl bromide (PC) 0.009 mM
Upper Limit	0.093	0.686	4.2%	59.0%	0.170	101.4%	103.7%	94.3%	77.5%	48.9%	32.1%	19.9%
Lower Limit	0.067	0.318	-1.6%	44.7%	0.009	93.8%	87.3%	65.1%	30.4%	12.9%	1.5%	-3.8%

PDA fluorescence QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Glutaralde-hyde (PC) 0.0200 mM	Glutaralde-hyde (PC) 0.0100 mM	Glutaralde-hyde (PC) 0.0050 mM	Glutaralde-hyde (PC) 0.0025 mM	Glutaralde-hyde (PC) 0.0013 mM	Glutaralde-hyde (PC) 0.0006 mM	Glutaralde-hyde (PC) 0.0003 mM
Upper Limit	41.4	316.6	7.6%	63.1%	0.008	106.5%	95.6%	65.1%	38.3%	22.0%	21.8%	26.4%
Lower Limit	16.5	15.3	-2.7%	34.8%	0.003	83.7%	58.0%	19.3%	0.7%	-5.2%	-14.5%	-20.8%

Abbreviations: COV = coefficient of variation; IC_{50} = half-maximal inhibitory concentration; NBT = 4-nitrobenzenthiol; NC = negative control; PC = positive control; PDA = pyridoxylamine; QC = quality control.

6.2.1.3. DoD

DoD completed the NBT pre-validation plates in early August 2019 and an initial PDA set shortly thereafter. However, analysis by NIST revealed what was ultimately determined to be excess PDA signal, causing overflow and a resulting “rightward” bias in the plates. DoD ultimately titrated the PDA concentration to half the originally prescribed PDA concentration (1 mM) to 0.5 mM. The PDA QC pre-validation plates were completed at the end of March 2021 due to the COVID-19 shutdown. DoD QC parameters used for the validation study are shown in **Table 9**.

Table 9. DoD baseline QC parameters

NBT absorbance QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Benzyl bromide (PC) 0.600 mM	Benzyl bromide (PC) 0.300 mM	Benzyl bromide (PC) 0.150 mM	Benzyl bromide (PC) 0.075 mM	Benzyl bromide (PC) 0.038 mM	Benzyl bromide (PC) 0.019 mM	Benzyl bromide (PC) 0.009 mM
Upper Limit	0.082	1.116	5.2%	56.8%	0.097	100.3%	100.4%	92.9%	73.3%	45.9%	25.9%	17.5%
Lower Limit	0.062	0.365	-2.3%	49.7%	0.066	96.9%	93.1%	72.4%	38.8%	18.0%	8.2%	-0.7%

PDA fluorescence QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Glutaralde-hyde (PC) 0.0200 mM	Glutaralde-hyde (PC) 0.0100 mM	Glutaralde-hyde (PC) 0.0050 mM	Glutaralde-hyde (PC) 0.0025 mM	Glutaralde-hyde (PC) 0.0013 mM	Glutaralde-hyde (PC) 0.0006 mM	Glutaralde-hyde (PC) 0.0003 mM
Upper Limit	3460104.9	18550819.3	14.5%	108.2%	0.015	111.7%	109.0%	117.9%	162.3%	179.8%	137.6%	82.7%
Lower Limit	242060.1	3268105.4	-6.8%	7.6%	-0.015	81.2%	83.8%	68.6%	3.1%	-51.1%	-59.1%	-35.1%

Abbreviations: COV = coefficient of variation; IC_{50} = half-maximal inhibitory concentration; NBT = 4-nitrobenzenthiol; NC = negative control; PC = positive control; PDA = pyridoxylamine; QC = quality control.

6.2.1.4. BRT

BRT joined the validation study in 2021, with completion of the baseline parameter QC plates in December 2021. BRT and DoD were utilizing the same model of plate reader, and therefore pre-emptively titrated the PDA concentration for the pre-validation plates. **Table 10** shows BRT's baseline QC parameters.

Table 10. BRT baseline QC parameters

NBT absorbance QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Benzyl bromide (PC) 0.600 mM	Benzyl bromide (PC) 0.300 mM	Benzyl bromide (PC) 0.150 mM	Benzyl bromide (PC) 0.075 mM	Benzyl bromide (PC) 0.038 mM	Benzyl bromide (PC) 0.019 mM	Benzyl bromide (PC) 0.009 mM
Upper Limit	0.099	0.886	6.8%	55.6%	0.147	100.6%	99.3%	93.0%	55.5%	34.0%	22.6%	12.0%
Lower Limit	0.057	0.386	-0.9%	49.9%	0.062	94.8%	86.5%	47.4%	29.3%	11.5%	1.0%	3.5%

PDA fluorescence QC report data

	NC/PC Blank	NC	NC COV	PC Midpoint	PC Forecast IC_{50} (mM)	Glutaralde-hyde (PC) 0.0200 mM	Glutaralde-hyde (PC) 0.0100 mM	Glutaralde-hyde (PC) 0.0050 mM	Glutaralde-hyde (PC) 0.0025 mM	Glutaralde-hyde (PC) 0.0013 mM	Glutaralde-hyde (PC) 0.0006 mM	Glutaralde-hyde (PC) 0.0003 mM
Upper Limit	1277827.0	10919037.1	12.8%	65.0%	0.008	97.5%	100.8%	120.5%	57.4%	50.6%	34.5%	34.8%
Lower Limit	597433.3	3685662.5	-0.1%	39.2%	0.000	92.3%	79.3%	10.3%	13.1%	-3.6%	-1.1%	-12.8%

Abbreviations: COV = coefficient of variation; IC_{50} = half-maximal inhibitory concentration; NC = negative control; NBT = 4-nitrobenzenthiol; PC = positive control; PDA = pyridoxylamine; QC = quality control.

7. Within-laboratory Reproducibility

The assessment of within-laboratory reproducibility (WLR) is based upon 12 of the PS reference chemicals (**Table 11**) (OECD, 2019). Each chemical required three QTs. Each QT is comprised of "at least two qualified independent repetitions". If the first two repetitions were not concordant, a third repetition was required to determine the outcome. WLR was calculated as an average of the reproducibility of the NBT and the PDA, which are then combined to make a prediction.

Table 11. Test chemicals for within-laboratory reproducibility

Chemical Name	CAS No.	LLNA EC3 (%)	In Vivo Prediction	DPRA Results	ADRA Results
Lauryl gallate	1166-52-5	0.3	Strong sensitizer	Pos	Pos
Chloramine trihydrate	127-65-1	0.4	Strong sensitizer	Pos	Pos
4-(Methylamino) phenol, hemisulfate	55-55-0	0.8	Strong sensitizer	Pos	Pos
2-Mercaptobenzothiazole	149-30-4	1.7	Moderate sensitizer	Pos	Pos
Benzyl salicylate	118-58-1	2.9	Moderate sensitizer	Pos/Neg	Neg
Cinnamaldehyde	14371-10-9 (104-55-2)	3	Moderate sensitizer	Pos	Pos
Imidazolidinyl urea	39236-46-9	24	Moderate sensitizer	Pos	Pos
Ethyl acrylate	140-88-5	28	Weak sensitizer	Pos	Pos
Salicylic acid	69-72-7	-	Nonsensitizer	Pos/Neg	Neg
Benzyl alcohol	100-51-6	-	Nonsensitizer	Pos/Neg	Neg
Glycerol	56-81-5	-	Nonsensitizer	Neg	Neg
Isopropanol	67-63-0	-	Nonsensitizer	Neg	Neg

Abbreviations: ADRA = amino acid derivative reactivity assay; CAS = Chemical Abstracts Service; DPRA = direct peptide reactivity assay; EC3 = dose (weight %) that gives a stimulation index of 3, positive threshold; LLNA = local lymph node assay.

7.1. Lead Lab (CPSC/NIST)

CPSC/NIST conducted 37 NBT assays and 33 PDA assays, with a total of three QC sheets needed for the PDA assays due to plate and PDA stock lot-to-lot variation. Of these assays, two NBT and six PDA plates failed the forecast IC_{50} criterion for the PC. All other plates passed and were included in the assessment of performance for the WLR. There were no failures on an individual chemical basis (due to lack of adequate replicates). One chemical was inconclusive due to interference with PDA, so no definitive call could be made (NBT was also negative).

For the 12 reference chemicals, CPSC/NIST had a WLR (calculated by QT) of 94%. For the individual probes, the WLR for NBT was 92% and for PDA was 93%. **Table 12** shows the individual run call for each probe and TC, with the EASA call for each TC. The only chemical misclassified was the weak sensitizer ethyl acrylate. Salicylic acid (a nonsensitizer) was inconclusive due to the fact that the test article was diluted in the PDA assay in an attempt to avoid fluorescence interference. While each individual PDA assay resulted in a negative call, results must be taken with caution as the chemical concentration was lower than that prescribed by the protocol. Thus, an inconclusive result was assigned.

Table 12. CPSC/NIST WLR results by QT

Chemical Name	LLNA Potency	QT1 - NBT	QT1 - PDA	QT1 Call	QT2 - NBT	QT2 - PDA	QT2 Call	QT3 - NBT	QT3 - PDA	QT3 Call	Overall Call
Lauryl gallate	Strong	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Pos-Pos	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Chloramine trihydrate	Strong	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
4-(Methylamino) phenol, hemisulfate	Strong	Pos-Pos-N/A	N/A-N/A-N/A	Pos	Pos-Pos-N/A	N/A-N/A-N/A	Pos	Pos-Pos-N/A	N/A-N/A-N/A	Pos	Pos
2-Mercapto-benzothiazole	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Benzyl salicylate	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Pos
Cinnamaldehyde	Moderate	Neg-Pos-Neg	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Imidazolidinyl urea	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	N/A-N/A-N/A	Inc	Pos
Ethyl acrylate	Weak	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Pos-Neg-Neg	Neg	Neg-Neg-N/A	Pos-Neg-Neg	Neg	Neg
Salicylic acid	Nonsensitizer	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Neg-Neg-N/A	N/A-N/A-N/A	Inc	Inc
Benzyl alcohol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Pos-Pos	Neg-Pos-Pos	Pos	Neg
Glycerol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Pos-Neg	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Isopropanol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg

Abbreviations: LLNA = local lymph node assay; N/A = not applicable; NBT = 4-nitrobenzenthiol; Neg = negative; PDA = pyridoxylamine; Pos = positive; QT = qualified test.

7.2. FDA

FDA conducted 19 NBT assays and 25 PDA assays. Of these assays, one NBT and three PDA plates failed the forecast IC_{50} criterion for the PC. All other plates passed and were included in the assessment of performance for the WLR. There were five failures on an individual chemical basis (due to lack of adequate replicate wells for that chemical on a plate).

For the 12 reference chemicals, FDA had WLR (calculated by QT) of 100%. For the individual probes, the WLR for NBT was 100% and for PDA was 100%. **Table 13** shows the individual run call for each probe and TC, with the final overall call for each TC. Two chemicals were misclassified, the weak sensitizer ethyl acrylate and the nonsensitizer benzyl alcohol.

Table 13. FDA WLR results by QT

Chemical Name	LLNA Potency	QT1 - NBT	QT1 - PDA	QT1 Call	QT2 - NBT	QT2 - PDA	QT2 Call	QT3 - NBT	QT3 - PDA	QT3 Call	Overall Call
Lauryl gallate	Strong	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Pos
Chloramine trihydrate	Strong	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
4-(Methylamino) phenol, hemisulfate	Strong	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Pos-Neg	Pos	Pos
2-Mercapto-benzothiazole	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Pos
Benzyl salicylate	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Pos
Cinnamaldehyde	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Neg-Pos-Pos	Pos-Pos-N/A	Pos	Pos
Imidazolidinyl urea	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Pos-Neg	Pos-Pos-N/A	Pos	Pos
Ethyl acrylate	Weak	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Salicylic acid	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Benzyl alcohol	Nonsensitizer	Pos-Neg-Pos	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos
Glycerol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Isopropanol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg

Abbreviations: LLNA = local lymph node assay; N/A = not applicable; NBT = 4-nitrobenzenthiol; Neg = negative; PDA = pyridoxylamine; Pos = positive; QT = qualified test.

7.3. DoD

DoD conducted 27 NBT assays and 32 PDA assays. Of these assays, four NBT plates failed the forecast IC_{50} criterion for the PC and two PDA plates failed the NC COV criteria. All other plates passed and were included in the assessment of performance for the WLR. There were no failures on an individual chemical basis. One chemical was inconclusive due to interference with PDA, so no definitive call could be made (NBT was also negative).

For the 12 reference chemicals, DoD had a WLR (calculated by QT) of 100%. For the individual probes, the WLR for NBT was 92% and for PDA was 100%.

Table 14 shows the individual run call for each probe and TC, with the final overall call for each TC. The only chemical misclassified was the weak sensitizer ethyl acrylate. Salicylic acid was inconclusive due to the fact that the test article was diluted in the PDA assay in an attempt to avoid fluorescence interference. While each individual PDA assay resulted in a negative call, results must be taken with caution as the chemical concentration was lower than that prescribed by the protocol. Thus, an inconclusive result was assigned.

Table 14. DoD WLR results by QT

Chemical Name	LLNA Potency	QT1 - NBT	QT1 - PDA	QT1 Call	QT2 - NBT	QT2 - PDA	QT2 Call	QT3 - NBT	QT3 - PDA	QT3 Call	Overall Call
Lauryl gallate	Strong	Neg-Neg-N/A	Neg-Pos-Pos	Pos	Pos-Neg-Pos	Pos-Neg-Pos	Pos	Neg-Pos-Pos	Pos-Pos-N/A	Pos	Pos
Chloramine trihydrate	Strong	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
4-(Methylamino) phenol, hemisulfate	Strong	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos
2-Mercapto-benzothiazole	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Benzyl salicylate	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Pos
Cinnamaldehyde	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Imidazolidinyl urea	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Pos
Ethyl acrylate	Weak	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Salicylic acid	Nonsensitizer	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Inc
Benzyl alcohol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Glycerol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Isopropanol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg

Abbreviations: LLNA = local lymph node assay; N/A = not applicable; NBT = 4-nitrobenzenthiol; Neg = negative; PDA = pyridoxylamine; Pos = positive; QT = qualified test.

7.4. BRT

BRT conducted 23 NBT assays and 34 PDA assays. All plates passed and were included in the assessment of performance for the WLR. There were no failures on an individual chemical basis. One chemical was inconclusive due to interference with PDA, so no definitive call could be made (NBT was also negative).

For the 12 reference chemicals, BRT had an overall WLR (calculated by QT) of 100%. For the individual probes, the WLR for NBT was 94% and for PDA was 94%. **Table 15** shows the individual run call for each probe and TC, with the final overall call for each TC. There were two chemicals misclassified, the weak sensitizer ethyl acrylate and the nonsensitizer benzyl alcohol. Salicylic acid was inconclusive due to the fact that the test article was diluted in the PDA assay in an attempt to avoid fluorescence interference. While each individual PDA assay resulted in a negative call, results must be taken with caution as the chemical concentration was lower than that prescribed by the protocol (2 mM). Thus, an inconclusive result was assigned.

Table 15. BRT WLR results by QT

Chemical Name	LLNA Potency	QT1 - NBT	QT1 - PDA	QT1 Call	QT2 - NBT	QT2 - PDA	QT2 Call	QT3 - NBT	QT3 - PDA	QT3 Call	Overall Call
Lauryl gallate	Strong	Pos-Pos-N/A	Pos-Neg-Pos	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Neg-Pos-Pos	Pos	Pos
Chloramine trihydrate	Strong	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
4-(Methylamino) phenol, hemisulfate	Strong	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos-Pos-N/A	Inc-Inc-N/A	Pos	Pos
2-Mercapto-benzothiazole	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Benzyl salicylate	Moderate	Pos-Neg-Pos	Pos-Pos-N/A	Pos	Neg-Pos-Neg	Pos-Pos-N/A	Pos	Neg-Neg-Pos	Pos-Pos-N/A	Pos	Pos
Cinnamaldehyde	Moderate	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos-Pos-N/A	Pos-Pos-N/A	Pos	Pos
Imidazolidinyl urea	Moderate	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Neg-Pos-Neg	Pos-Pos-N/A	Pos	Neg-Neg-N/A	Pos-Pos-N/A	Pos	Pos
Ethyl acrylate	Weak	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Pos-Neg-Neg	Neg-Neg-N/A	Neg	Neg-Pos-Neg	Neg-Pos-Neg	Neg	Neg
Salicylic acid	Nonsensitizer	Neg-Pos-Neg	Inc-Inc-N/A	Inc	Pos-Neg-Pos	Inc-Inc-N/A	Pos	Neg-Neg-N/A	Inc-Inc-N/A	Inc	Inc
Benzyl alcohol	Nonsensitizer	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos-Pos-N/A	Neg-Neg-N/A	Pos	Pos
Glycerol	Nonsensitizer	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg-Neg-N/A	Neg-Neg-N/A	Neg	Neg
Isopropanol	Nonsensitizer	Pos-Neg-Neg	Pos-Neg-Neg	Neg	Neg-Pos-Neg	Neg-Pos-Neg	Neg	Neg-Pos-Neg	Neg-Pos-Neg	Neg	Neg

Abbreviations: LLNA = local lymph node assay; N/A = not applicable; NBT = 4-nitrobenzenthiol; Neg = negative; PDA = pyridoxylamine; Pos = positive; QT = qualified test.

7.5. Overall Within-laboratory Reproducibility

The average WLR for EASA calls from all four laboratories by QT was 98%, with a WLR of 94% at CPSC/NIST and 100% at the other three participating labs. The probe-specific WLRs were 95% and 97% for NBT and PDA, respectively. A summary of individual QT calls by lab is found in **Table 16.** Three of the four labs had inconclusive calls on salicylic acid. For those three labs, the test article was diluted below 2 mM concentration specified in the protocol in an attempt to avoid the strong fluorescent signal that interfered with the PDA probe. As such, it was not possible to make a definitive negative EASA call for these labs, as a negative required that both NBT and PDA be negative. The fourth lab, FDA, which was able to make a conclusive call, had much lower fluorescence signal on their plate reader, which limited the overflow interference and allowed for a conclusive test. A similar issue was found with Metol, where the same three labs were not able to make conclusive calls for the test article with the PDA probe. However, as it was positive in the NBT, it was correctly classified by the EASA. Two labs (FDA and BRT) also misclassified benzyl alcohol as positive, with a third having a positive result in one of their QTs (CPSC/NIST). These mixed results are consistent with the mixed results found in the DPRA validation report as well (EURL ECVAM, 2012). Additionally, the NBT probe did not correctly classify benzyl salicylate or imidazolidynl urea as positive for any of the labs, but the PDA did, which still resulted in a positive EASA call.

Table 16. Replicate EASA QT calls per lab

Chemical Name	CPSC/NIST	FDA	DoD	BRT
Lauryl gallate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Chloramine trihydrate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
4-(Methylamino) phenol, hemisulfate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
2-Mercapto-benzothiazole	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Benzyl salicylate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Cinnamaldehyde	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Imidazolidinyl urea	Pos-Pos-Inc	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Ethyl acrylate	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg
Salicylic acid	Inc-Inc-Inc	Neg-Neg-Neg	Inc-Inc-Inc	Inc-Pos-Inc
Benzyl alcohol	Neg-Neg-Pos	Pos-Pos-Pos	Neg-Neg-Neg	Pos-Pos-Pos
Glycerol	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg
Isopropanol	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg

Abbreviations: BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DoD = U.S. Department of Defense; FDA = U.S. Food and Drug Administration; Inc = inconclusive; Neg = negative; Pos = positive.

8. Between-laboratory Reproducibility

The assessment of between-laboratory reproducibility (BLR) is based upon all 20 reference chemicals within the PS (OECD, 2019). Eight of these chemicals (**Table 17**) were tested in 1 QT, while a consensus result based on concordance of each QT was produced for each of the 12 chemicals tested in the WLR phase.

Table 17. Test chemicals for between-laboratory reproducibility

Chemical Name	CAS No.	LLNA EC3 (%)	In Vivo Prediction	DPRA Results	ADRA Results
p-Benzoquinone	106-51-4	0.0099	Extreme sensitizer	Pos	Pos
Dihydroeugenol	2785-87-7	6.8	Moderate sensitizer	Pos/Neg	Pos/Neg
Palmitoyl chloride	112-67-4	8.8	Moderate sensitizer	Pos	Pos
Farnesal	19317-11-4 (502-67-0)	12	Weak sensitizer	Pos	Pos
Benzyl cinnamate	103-41-3	18	Weak sensitizer	Neg	Neg
Dimethyl isophthalate	1459-93-4	-	Nonsensitizer	Neg	Neg
Methyl salicylate	119-36-8	-	Nonsensitizer	Pos/Neg	Neg
4-Aminobenzoic acid	150-13-0	-	Nonsensitizer	Neg	Neg

Abbreviations: ADRA = amino acid derivative reactivity assay; CAS = Chemical Abstracts Service; DPRA = direct peptide reactivity assay; EC3 = dose (weight %) that gives a stimulation index of 3, positive threshold; LLNA = local lymph node assay.

Of the 20 TCs, three were discordant across the different testing facilities: salicylic acid, benzyl alcohol, and dihydroeugenol. Thus, BLR was 85% (**Table 18**). For salicylic acid the outcome was inconclusive for all but FDA, due to the need to dilute the test article to resolve fluorescence interfere with the PDA probe. The EASA results for benzyl alcohol were mixed, with two negatives and two positives for the labs. These results are similar to the mixed results for DPRA from the validation study (EURL ECVAM, 2012). The pro-hapten dihydroeugenol also had mixed results: three labs had a negative result, while FDA had a positive result. These mixed results were consistent with both the DPRA and ADRA validation studies and for other in vitro skin sensitization assays, e.g., negative in the human cell line activation test but positive in KeratinoSens (OECD, 2025b). In human tests, it is also negative (Strickland et al., 2023).

Table 18. Individual laboratory results for BLR

Chemical Name	LLNA Reference	CPSC/NIST	FDA	DoD	BRT
Lauryl gallate	Strong	Pos	Pos	Pos	Pos
Chloramine T trihydrate	Strong	Pos	Pos	Pos	Pos
Metol (4-methylamino phenol)	Strong	Pos	Pos	Pos	Pos
2-Mercaptobenzothiazole	Moderate	Pos	Pos	Pos	Pos
Benzyl salicylate	Moderate	Pos	Pos	Pos	Pos
Cinnamaldehyde	Moderate	Pos	Pos	Pos	Pos
Imidazolidynl urea	Moderate	Pos	Pos	Pos	Pos
Ethyl acrylate	Weak	Neg	Neg	Neg	Neg
Salicylic acid	Nonsensitizer	Inc	Neg	Inc	Inc
Benzyl alcohol	Nonsensitizer	Neg	Pos	Neg	Pos
Glycerol	Nonsensitizer	Neg	Neg	Neg	Neg
Isopropanol	Nonsensitizer	Neg	Neg	Neg	Neg
p-Benzoquinone	Extreme	Pos	Pos	Pos	Pos
Dihydroeugenol	Moderate	Neg	Pos	Neg	Neg
Palmitoyl chloride	Moderate	Pos	Pos	Pos	Pos
Farnesal	Weak	Pos	Pos	Pos	Pos
Benzyl cinnamate	Weak	Pos	Pos	Pos	Pos
Dimethyl isophthalate	Nonsensitizer	Neg	Neg	Neg	Neg
Methyl salicylate	Nonsensitizer	Pos	Pos	Pos	Pos
4-Aminobenzoic acid	Nonsensitizer	Neg	Neg	Neg	Neg

Abbreviations: BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DoD = U.S. Department of Defense; FDA = U.S. Food and Drug Administration; Inc = inconclusive; Neg = negative; Pos = positive.
Gray highlights indicate where individual calls were discordant between labs.

9. Predictive Capacity

For the 20 reference chemicals, predictive capacity was determined by calculating sensitivity, specificity, and accuracy in reference to existing LLNA data (OECD, 2019). The performance against human data was also assessed for 13 reference chemicals. A summary of the four

participating lab test outcomes for the 20 TCs is shown in **Table 19**, with a summary of consensus predictions and reference data in **Table 20**. Consensus was determined by concordance across three tests for the 12 TCs that were assessed with three QTs. Performance for each lab is shown in **Table 20**. Accuracy against the LLNA was 84% for CPSC/NIST, 85% for FDA, 84% for DoD, and 79% for BRT, with a cumulative accuracy of 83%. Sensitivity for CPSC/NIST was 85%, 92% for FDA, and 85% for DoD and BRT, while specificity was 83% for CPSC/NIST, 71% for FDA, 83% for DoD, and 67% for BRT, **Table 21**. The cumulative sensitivity was 87%, and specificity was 76%. When compared against human data, the accuracy for CPSC/NIST and DoD was 77%, 79% for FDA, and 85% for BRT, **Table 22**. Cumulative sensitivity and specificity were 81% and 76%, with cumulative accuracy of 79% for the labs against human data.

Table 19. Results from each lab by QT

Chemical Name	CPSC/NIST	FDA	DoD	BRT
Lauryl gallate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Chloramine trihydrate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
4-(Methylamino) phenol, hemisulfate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
2-Mercapto-benzothiazole	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Benzyl salicylate	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Cinnamaldehyde	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Imidazolidinyl urea	Pos-Pos-Inc	Pos-Pos-Pos	Pos-Pos-Pos	Pos-Pos-Pos
Ethyl acrylate	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg
Salicylic acid	Inc-Inc-Inc	Neg-Neg-Neg	Inc-Inc-Inc	Inc-Pos-Inc
Benzyl alcohol	Neg-Neg-Pos	Pos-Pos-Pos	Neg-Neg-Neg	Pos-Pos-Pos
Glycerol	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg
Isopropanol	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg	Neg-Neg-Neg
p-Benzoquinone	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A
Dihydroeugenol	Neg-N/A-N/A	Pos-N/A-N/A	Neg-N/A-N/A	Neg-N/A-N/A
Palmitoyl chloride	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A
Farnesal	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A
Benzyl cinnamate	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A
Dimethyl isophthalate	Neg-N/A-N/A	Neg-N/A-N/A	Neg-N/A-N/A	Neg-N/A-N/A
Methyl salicylate	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A	Pos-N/A-N/A
4-Aminobenzoic acid	Neg-N/A-N/A	Neg-N/A-N/A	Neg-N/A-N/A	Neg-N/A-N/A

Abbreviations: BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology;
DoD = U.S. Department of Defense; FDA = U.S. Food and Drug Administration; Inc = inconclusive; Neg = negative; Pos = positive; QT = qualified test.

Table 20. Consensus results from each lab (2 out of 3) and reference data

Chemical Name	LLNA Ref.	Human Ref.*	DPRA Ref.	ADRA Ref.	CPSC/ NIST	FDA	DoD	BRT
Lauryl gallate	Strong	Pos (1)	Pos	Pos	Pos	Pos	Pos	Pos
Chloramine T trihydrate	Strong	N/A	Pos	Pos	Pos	Pos	Pos	Pos
Metol (4-methylamino phenol)	Strong	Pos (1)	Pos	Pos	Pos	Pos	Pos	Pos
2-Mercaptobenzothiazole	Moderate	Pos (2)	Pos	Pos	Pos	Pos	Pos	Pos
Benzyl salicylate	Moderate	Pos/ Neg (3)	Pos/ Neg	Neg	Pos	Pos	Pos	Pos
Cinnamaldehyde	Moderate	Pos (2)	Pos	Pos	Pos	Pos	Pos	Pos
Imidazolidynl urea	Moderate	Pos (2)	Pos	Pos	Pos	Pos	Pos	Pos
Ethyl acrylate	Weak	Pos (2)	Pos	Pos	Neg	Neg	Neg	Neg
Salicylic acid	Nonsensitizer	Neg (3)	Pos/ Neg	Neg	Inc	Neg	Inc	Inc
Benzyl alcohol	Nonsensitizer	Pos (3)	Pos/ Neg	Neg	Neg	Pos	Neg	Pos
Glycerol	Nonsensitizer	Neg (3)	Neg	Neg	Neg	Neg	Neg	Neg
Isopropanol	Nonsensitizer	Neg (1)	Neg	Neg	Neg	Neg	Neg	Neg
p-Benzoquinone	Extreme	N/A	Pos	Pos	Pos	Pos	Pos	Pos
Dihydroeugenol	Moderate	Neg (3)	Pos/ Neg	Pos/ Neg	Neg	Pos	Neg	Neg
Palmitoyl chloride	Moderate	N/A	Pos	Pos	Pos	Pos	Pos	Pos
Farnesal	Weak	N/A	Pos	Pos	Pos	Pos	Pos	Pos
Benzyl cinnamate	Weak	Pos (1)	Neg	Neg	Pos	Pos	Pos	Pos
Dimethyl isophthalate	Nonsensitizer	N/A	Neg	Neg	Neg	Neg	Neg	Neg
Methyl salicylate	Nonsensitizer	Neg (3)	Pos/ Neg	Neg	Pos	Pos	Pos	Pos
4-Aminobenzoic acid	Nonsensitizer	Neg (3)	Neg	Neg	Neg	Neg	Neg	Neg

*Numbers in parentheses indicate source of human reference data: 1 = Basketter et al., 2015; 2 = TG 497 human predictive patch test database; 3 = NICEATM human predictive patch test database.
Abbreviations: ADRA = amino acid derivative reactivity assay; BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DoD = U.S. Department of Defense; DPRA = direct peptide reactivity assay; FDA = U.S. Food and Drug Administration; Inc = inconclusive; LLNA = local lymph node assay; Neg = negative; Pos = positive; Ref = reference.

Table 21. Predictive capacity for each participating lab against LLNA for 20 PS chemicals

Overview of results

Reference Result (LLNA)	Cumulative Pos	Cumulative Neg	CPSC/NIST Pos	CPSC/NIST Neg	FDA Pos	FDA Neg	DoD Pos	DoD Neg	BRT Pos	BRT Neg
Pos (N = 13)	45	7	11	2	12	1	11	2	11	2
Neg (N = 7)	6	19	1	5	1	5	1	5	2	4

Predictivity Statistics

Statistic	Cumulative	CPSC/NIST	FDA	DoD	BRT
Sensitivity	87%	85%	92%	85%	85%
Specificity	76%	83%	71%	83%	67%
Accuracy	83%	84%	85%	84%	79%

Abbreviations: BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DoD = U.S. Department of Defense; FDA = U.S. Food and Drug Administration; LLNA = local lymph node assay; Neg = negative; Pos = positive.

Table 22. Predictive capacity for each participating lab against human data (13 chemicals)

Overview of results

Reference Result (LLNA)	Cumulative Pos	Cumulative Neg	CPSC/NIST Pos	CPSC/NIST Neg	FDA Pos	FDA Neg	DoD Pos	DoD Neg	BRT Pos	BRT Neg
Pos (N = 8)	26	6	6	2	7	1	6	2	7	1
Neg (N = 5)	5	16	1	4	2	4	1	4	1	4

Predictivity Statistics

Statistic	Cumulative	CPSC/NIST	FDA	DoD	BRT
Sensitivity	81%	75%	88%	75%	88%
Specificity	76%	80%	67%	80%	80%
Accuracy	79%	77%	79%	77%	85%

Abbreviations: BRT = Burleson Research Technologies, Inc.; CPSC/NIST = U.S. Consumer Product Safety Commission/National Institute of Standards and Technology; DoD = U.S. Department of Defense; FDA = U.S. Food and Drug Administration; LLNA = local lymph node assay; Neg = negative; Pos = positive.

Table 23 shows a comparison of the EASA hazard calls for each reference substance to hazard call from other reference tests. While the overall predictive capacity of the assay for the labs had a specificity below the cut-off of 80%, this likely is due to the applicability domain of the assay. The primary chemicals that were misclassified were the pro-hapten dihydroeugenol, the weak sensitizer ethyl acrylate, and the nonsensitizer methyl salicylate. Additionally, a final call on salicylic acid could not be made due to interference with the PDA probe, resulting in an inconclusive result. However, when compared to the results for these same chemicals for DPRA and ADRA, several also missed the call entirely or had mixed results similar to EASA (**Table 22**).

Table 23. Comparison of EASA calls to reference test calls

Test Chemical	Mechanism	EASA	LLNA	DPRA	ADRA	Potency
p-Benzoquinone	Michael acceptor	Pos	Pos	Pos	Pos	Extreme
Lauryl gallate	Pre-hapten Michael acceptor	Pos	Pos	Pos	Pos	Strong
Chloramine trihydrate	Acylation	Pos	Pos	Pos	Pos	Strong
4-(Methylamino) phenol, hemisulfate	Pre-hapten Michael acceptor	Pos	Pos	Pos	Pos	Strong
2-Mercaptobenzothiazole	SN2, acylation	Pos	Pos	Pos	Pos	Moderate
Benzyl salicylate	SN2, acylation	Pos	Pos	Pos/Neg	Neg	Moderate
Cinnamaldehyde*	Michael acceptor	Pos	Pos	Pos	Pos	Moderate
Imidazolidinyl urea	Acylation	Pos	Pos	Pos	Pos	Moderate
Dihydroeugenol*	Pro-hapten, SN2, Michael acceptor	Pos/Neg	Pos	Pos/Neg	Pos/Neg	Moderate
Palmitoyl chloride	Acylation	Pos	Pos	Pos	Pos	Moderate
Ethyl acrylate	Michael acceptor	Neg	Pos	Pos	Pos	Weak
Farnesal	Schiff base	Pos	Pos	Pos	Pos	Weak
Benzyl cinnamate	SN2, Acylation	Pos	Pos	Neg	Neg	Weak
Salicylic acid	Non-reactive	Neg/Inc	Neg	Pos/Neg	Neg	Nonsensitizer
Benzyl alcohol*	Non-reactive	Pos/Neg	Neg	Pos/Neg	Neg	Nonsensitizer
Glycerol	Non-reactive	Neg	Neg	Neg	Neg	Nonsensitizer
Isopropanol	Non-reactive	Neg	Neg	Neg	Neg	Nonsensitizer
Dimethyl isophthalate	Non-reactive	Neg	Neg	Neg	Neg	Nonsensitizer
Methyl salicylate	Non-reactive	Pos	Neg	Pos/Neg	Neg	Nonsensitizer
4-Aminobenzoic acid	Non-reactive	Neg	Neg	Neg	Neg	Nonsensitizer

*Chemicals replaced in 2022 Performance Standards document.
Abbreviations: ADRA = amino acid derivative reactivity assay; DPRA = direct peptide reactivity assay; EASA = electrophilic allergen screening assay; Inc = inconclusive; LLNA = local lymph node assay; Neg = negative; Pos = positive.
Yellow fill indicates where EASA results do not match LLNA reference results.
Green fill indicates where existing TG 442C methods do not match LLNA reference results.

In an expanded data set from the development of the assay (Petersen et al., 2022a) against in vivo animal data, the EASA had a sensitivity of 80%, specificity of 61%, and accuracy of 74%. Thus, the EASA may slightly overpredict skin sensitizers.

10. Applicability Domain

Due to the design and methodology of the EASA, there are inherent limitations to the type of chemicals that can be successfully assessed within the assay. Solubility and interference of TCs with the test system are primary limitations on the types of chemicals that can be tested. While the assay has been designed with multiple options for solvent type, highly non-polar TCs are unable to be adequately assessed. The EASA will also not adequately assess chemicals that react to light in a way that interferes with reading the signal from either of the two test probes. The assay does allow for dilution of the TC to a certain extent; however, negative tests must be considered inconclusive if they are tested below the 2 mM concentration specified by the protocol. As the EASA does not have any metabolic capacity, it is unable to reliably detect pro-haptens (such as dihydroeugenol) and has limited capacity to identify metal contact allergens.

Dimerization between the thiol (-SH) functional groups on NBT could also affect signal strength, causing a decrease in signal that is not due to interaction with the TC. However, this is a potential effect in all of the KE1-based test methods.

11. Quality Check

11.1. Chemical Handling and Issues

The NTP Chemical Repository handled acquisition and distribution of TCs and safety data sheets to participating labs. All documentation was sent to the chemical safety officer in a sealed envelope, to only be opened in case of an accident. None of the sealed envelopes had to be opened. The Chemical Repository distributed coded TCs to each of the chemical safety officers in February 2020 via MRIGlobal. Chemicals were shipped according to storage/safety instructions. The FDA completed all testing with this first lot of chemicals. It was found in February 2021 that some of the chemicals had expired, and a second set of newly coded TCs was shipped to the remaining three participating labs in June 2021. All chemicals except ethyl acrylate were noted as different lots from the same manufacturer(s). Participating laboratories were provided with instructions to achieve 10 mM stock concentrations for each chemical, but not the appropriate solvent to use.

11.2. Quality Control of Data

Test data were provided to NICEATM in the data calculator and QC chart Excel worksheets. Data for each reference chemical were reviewed and compiled by run in summary worksheets to ensure that an adequate number of runs had been performed to meet the number of QTs required for each test. At the completion of testing, NICEATM provided a summary of the run and QT test results to each lab for review to verify the tests (and results) that had been performed and submitted.

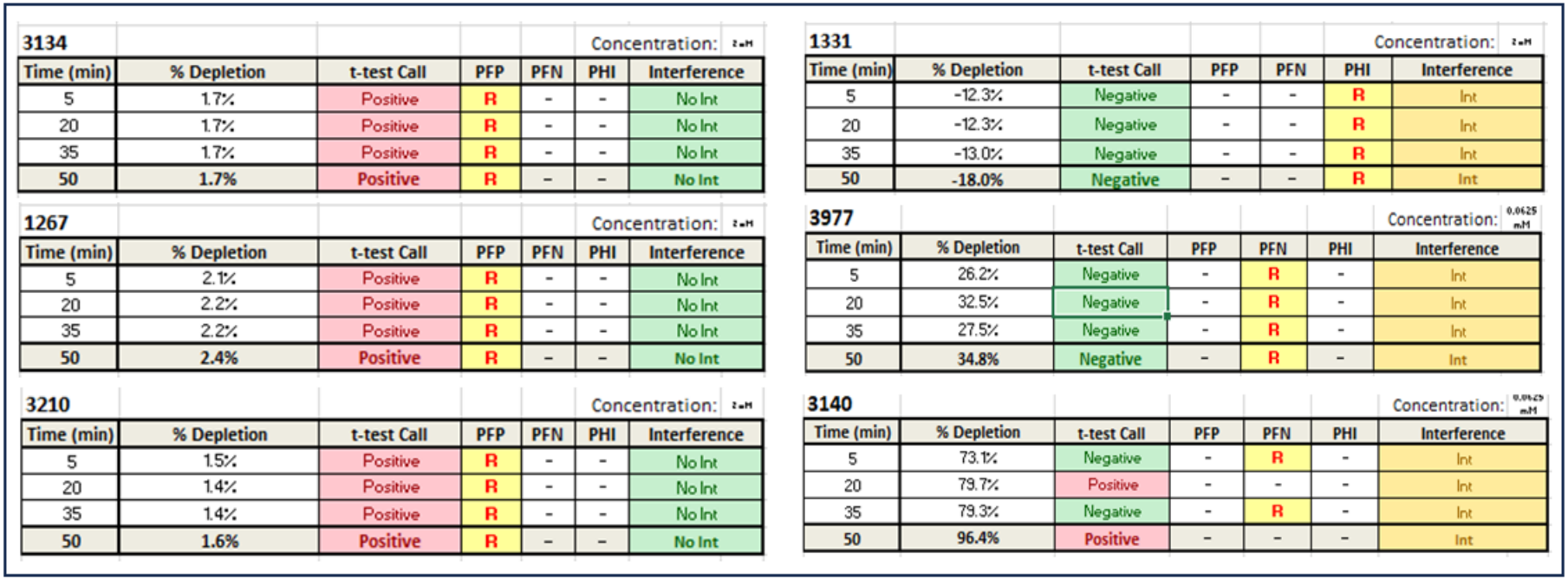
12. Overall Conclusions and Recommendations of the VMT

12.1. Conclusions

The objective of this validation study was to assess the performance (sensitivity, specificity, and accuracy) of the EASA test method, which may be able to predict skin sensitization hazard with higher throughput, greater accessibility, and lower cost than existing animal methods. The study also focused on optimizing the protocol and evaluating the transferability of the method, including assessing the WLR and BLR. Coded TCs were used, and testing was completed as defined in the TG 442C PS (OECD, 2019). TG 442C calls for at least 80% concordance with the PS for WLR, BLR, sensitivity, specificity, and accuracy, and a required explanation for failure to meet benchmarks or misclassify strong or extreme sensitizers. The WLR and BLR met both benchmarks easily, with 96% and 85%. WLRs by individual laboratory ranged from 83% to 100%. Probe-specific WLRs were 88% and 93% for NBT and PDA, respectively. For the performance of the assay, as evaluated against the LLNA, the overall sensitivity was 87%, specificity was 76%, and accuracy was 83%. The EASA specificity did not meet the 80% threshold. This was attributed to including several nonsensitizer TCs in the PS, which meant that where chemical interference occurred, i.e., auto-fluorescence against the PDA, a conclusive call for the test was not possible, such as with salicylic acid. This interference highlights the limitations of the assay, particularly with respect to the probe interference and a lack of metabolic capacity, which are similar to the limitations of the DPRA and ADRA.

An additional potential limitation is that the use of a statistical cut-off to make a call on sensitizing potential for the EASA, instead of a threshold cut-off as is used in the DPRA or ADRA, may have caused some chemicals to be listed as false-negatives or false-positives (see examples in Figure 7). When a statistical cut-off is used to determine that a chemical is a sensitizer, it requires over the assay's time course a statistically significant difference in the depletion values between the TC and NC. This difference is independent of the observed percent depletion for that TC. For example, **Figure 7** shows that chemical 3134 has a 1.7% depletion value but was determined to be positive. Alternatively, chemical 3977 was negative, despite having a depletion value of 34.8%. Although flagged as potentially false positive/negative, these results did not affect the final call for the EASA. It is possible that several chemicals may have been more accurately predicted if threshold cut-off values been for determining specificity instead of statistical cut-offs.

Figure 7. Screenshot examples of potential false positive or potential false negative calls from BRT data



3134 Concentration: 2 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	1.7%	Positive	R	-	-	No Int
20	1.7%	Positive	R	-	-	No Int
35	1.7%	Positive	R	-	-	No Int
50	1.7%	Positive	R	-	-	No Int

1267 Concentration: 2 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	2.1%	Positive	R	-	-	No Int
20	2.2%	Positive	R	-	-	No Int
35	2.2%	Positive	R	-	-	No Int
50	2.4%	Positive	R	-	-	No Int

3210 Concentration: 2 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	1.5%	Positive	R	-	-	No Int
20	1.4%	Positive	R	-	-	No Int
35	1.4%	Positive	R	-	-	No Int
50	1.6%	Positive	R	-	-	No Int

1331 Concentration: 2 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	-12.3%	Negative	-	-	R	Int
20	-12.3%	Negative	-	-	R	Int
35	-13.0%	Negative	-	-	R	Int
50	-18.0%	Negative	-	-	R	Int

3977 Concentration: 0.0625 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	26.2%	Negative	-	R	-	Int
20	32.5%	Negative	-	R	-	Int
35	27.5%	Negative	-	R	-	Int
50	34.8%	Negative	-	R	-	Int

3140 Concentration: 0.0625 mM

Time (min)	% Depletion	t-test Call	PFP	PFN	PHI	Interference
5	73.1%	Negative	-	R	-	Int
20	79.7%	Positive	-	-	-	Int
35	79.3%	Negative	-	R	-	Int
50	96.4%	Positive	-	-	-	Int

The number above each table refers to a coded TC for one of the testing facilities. The three tables on the left-side show instances where little depletion occurred (<5%) but was called a positive run. The three tables on the right-side show instances where a negative depletion occurred (TC 1331) or where the percent depletion would be expected to result in a positive call but is instead negative (TC 3977/TC 3140). The outcome for chemical 3140 in this run was still positive, but in the 5- and 35-minute reads, similar depletion values were considered negative. Notations of "R" in "PFN", "PFP", or "PHI" columns indicate scenarios in which additional testing might be needed to address specific regulatory information requirements.

Abbreviations: Int = interference; PFN = potential false negative; PFP = potential false positive; PHI = potential high interference; TC = test chemical.

12.2. Recommendations

This validation study was performed to enable assessment of the utility of the EASA method to address the MIE in the skin sensitization AOP and inform upon chemical protein reactivity. The information from this study may support future evaluation of the EASA as a "me-too" method within OECD TG 442C and whether it can be used as part of a DA or integrated approach to testing and assessment to support discrimination between skin sensitizers and nonsensitizers for the purpose of hazard classification and labeling and potency categorization according to the United Nations Globally Harmonized System of Classification and Labelling of Chemicals. Additional work could be performed as needed to support such an assessment. For example, the validation study was conducted in its entirety prior to the publication of the updated PS. Therefore, it may be beneficial for the lead lab and one other lab to conduct testing on the new chemicals in the updated PS list to determine if performance improves. Additionally, an assessment on the use of statistical cut-offs as compared to a threshold cut-off could be conducted. The method developers could also develop a decision process for how to handle potential false positive and false negative outcomes. With a progression toward utilizing defined approaches for determining hazard and potency, the EASA should also be evaluated for performance within these DAs.

13. References

Alinaghi, F., Bennike, N. H., Egeberg, A., Thyssen, J. P., & Johansen, J. D. (2019). Prevalence of contact allergy in the general population: A systematic review and meta-analysis. *Contact Dermatitis*, *80*(2), 77–85. https://doi.org/10.1111/cod.13119

Chipinda, I., Ajibola, R. O., Morakinyo, M. K., Ruwona, T. B., Simoyi, R. H., & Siegel, P. D. (2010). Rapid and simple kinetics screening assay for electrophilic dermal sensitizers using nitrobenzenethiol. *Chemical Research in Toxicology*, *23*(5), 918–925. https://doi.org/10.1021/tx100003w

Chipinda, I., Mbiya, W., Adigun, R. A., Morakinyo, M. K., Law, B. F., Simoyi, R. H., & Siegel, P. D. (2014). Pyridoxylamine reactivity kinetics as an amine based nucleophile for screening electrophilic dermal sensitizers. *Toxicology*, *315*, 102–109. https://doi.org/10.1016/j.tox.2013.11.009

Daniel, A. B., Strickland, J., Allen, D., Casati, S., Zuang, V., Barroso, J., Whelan, M., Régimbald-Krnel, M. J., Kojima, H., Nishikawa, A., Park, H.-K., Lee, J. K., Kim, T. S., Delgado, I., Rios, L., Yang, Y., Wang, G., & Kleinstreuer, N. (2018). International regulatory requirements for skin sensitization testing. *Regulatory Toxicology and Pharmacology: RTP*, *95*, 52–65. https://doi.org/10.1016/j.yrtph.2018.03.003

EPA. (2018). *Interim Science Policy: Use of Alternative Approaches for Skin Sensitization as a Replacement for Laboratory Animal Testing (Draft for Public Comment)*. U.S. Environmental Protection Agency.

EPA. (2016). *Frank R. Lautenberg Chemical safety for the 21st Century Act* (Public Law 114-182). U.S. Environmental Protection Agency. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/frank-r-lautenberg-chemical-safety-21st-century-act

EURL ECVAM. (2012). *Direct Peptide Reactivity Assay (DPRA) ECVAM Validation Study Report.* https://tsar.jrc.ec.europa.eu/test-method/tm2009-06

European Union. (2006). *Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC*. European Union. http://data.europa.eu/eli/reg/2006/1907/oj/eng

European Union. (2009). *Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products*. European Union. http://data.europa.eu/eli/reg/2009/1223/2019-08-13/eng

Lopachin, R. M., Gavin, T., Decaprio, A., & Barber, D. S. (2012). Application of the Hard and Soft, Acids and Bases (HSAB) theory to toxicant—Target interactions. *Chemical Research in Toxicology*, *25*(2), 239–251. https://doi.org/10.1021/tx2003257

Nel, A. E., & Malloy, T. F. (2017). Policy reforms to update chemical safety testing. *Science*, *355*(6329), 1016–1018. https://doi.org/10.1126/science.aak9919

OECD. (2010). *OECD guideline for the testing of chemicals. Skin Sensitization: Local Lymph Node Assay* (OECD TG 429). Organization for Economic Co-Operation and Development, Paris, France.

OECD. (2012). *The adverse outcome pathway for skin sensitisation initiated by covalent binding to proteins* (Document no. ENV/JM/MONO(2016)29; Series on Testing and Assessment). Organization for Economic Co-Operation and Development, Paris, France.

OECD. (2018). *OECD guideline for the testing of chemicals. Skin Sensitization. Local Lymph Node Assay: BrdU-ELISA or –FCM* (OECD TG 442B). Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-442b-skin-sensitization_9789264090996-en

OECD. (2019). *OECD Performance Standards for the assessment of proposed similar or modified in vitro skin sensitisation DPRA and ADRA test methods.* (No. 303; Series on Testing and Assessment). Organisation for Economic Co-operation and Development. https://www.oecd.org/chemicalsafety/testing/performance-standards.htm

OECD. (2025a). *OECD guideline for the testing of chemicals. In Vitro Skin Sensitisation: ARE-Nrf2 Luciferase Test Method* (OECD TG 442D). Organisation for Economic Co-

operation and Development. https://www.oecd-ilibrary.org/environment/test-no-442d-in-vitro-skin-sensitisation_9789264229822-en

OECD. (2025b). *OECD Guideline for the testing of chemicals. Defined Approaches on Skin Sensitisation* (OECD TG 497). Organisation for Economic Co-operation and Development. https://www.oecd.org/chemicalsafety/guideline-no-497-defined-approaches-on-skin-sensitisation-b92879a4-en.htm

OECD. (2025c). *OECD guideline for the testing of chemicals. In Chemico Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins* (OECD 442C). Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-442c-in-chemico-skin-sensitisation_9789264229709-en

OECD. (2025d). *OECD guideline for the testing of chemicals. In Vitro Skin Sensitisation: In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation* (OECD TG 442E). Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-442e-in-vitro-skin-sensitisation_9789264264359-en

Olusegun, A.O. & Martincigh, B. S. (2021). Allergic contact dermatitis: A significant environmental and occupational skin disease. *International Journal of Dermatology*, *60*(9), 1082–1091. https://doi.org/10.1111/ijd.15502

Petersen, E. J., Elliott, J. T., Gordon, J., Kleinstreuer, N. C., Reinke, E., Roesslein, M., & Toman, B. (2022). Technical framework for enabling high-quality measurements in new approach methodologies (NAMs). *ALTEX*. https://doi.org/10.14573/altex.2205081

Petersen, E. J., Uhl, R., Toman, B., Elliott, J. T., Strickland, J., Truax, J., & Gordon, J. (2022). Development of a 96-Well Electrophilic Allergen Screening Assay for Skin Sensitization Using a Measurement Science Approach. *Toxics*, *10*(5), 257. https://doi.org/10.3390/toxics10050257

Strickland, J., Abedini, J., Allen, D. G., Gordon, J., Hull, V., Kleinstreuer, N. C., Ko, H.-S., Matheson, J., Thierse, H.-J., Truax, J., Vanselow, J. T., & Herzler, M. (2023). A database of human predictive patch test data for skin sensitization. *Archives of Toxicology*, *97*, 2825–2837. https://doi.org/10.1007/s00204-023-03530-3

Appendix A: Chemical Coding Information

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix B: Protocol

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix C: NBT Data Calculator

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix D: PDA Data Calculator

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix E: EASA QC Chart

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix F: EASA Validation Study Worksheets

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix G: Lab Data

https://doi.org/10.22427/NICEATM-DATA-NICEATM-3

Appendix H: Peer Review of the Draft Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers

Peer review of the draft "Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers" was conducted by letter review by six scientific experts in July through November 2024. The selection of experts and conduct of the peer review were in accordance with established practice, Department of Health and Human Services peer review policies[1], and the Office of Management and Budget's Final Information Quality Bulletin for Peer Review[2]. The peer reviewers served as independent scientists, not as representatives of any institution, company, or government agency. In this capacity, the reviewers had two major responsibilities: (1) Peer review the draft "Report on the Validation of the Electrophilic Allergen Screening Assay (EASA) for the Identification of Potential Skin Sensitizers" and (2) comment on whether the draft document was clearly written and objectively presented.

The peer reviewers agreed that the report was well written, sufficiently detailed, and easy to follow for understanding how the studies were conducted. They concurred that the report documented adequate performance using appropriate standards, including 20 reference chemicals from the update list, and that the authors achieved the objective of optimizing the test method protocol, assessing inter- and intra-laboratory reproducibility, and confirming accuracy for the classification of hazard. They also noted that the report presented a substantial amount of evidence supporting the advantages of this method in terms of costs, run time, throughput potential, and animal welfare benefits and was transparent regarding difficulties encountered.

The peer reviewers asked for additional text on the need for hypersensitivity testing, the differences between rodent and human responses to allergens in the context of interpreting the local lymph node assay data used as a reference, and the need for the EASA assay. They also requested further explanation on how participating laboratories were identified, the definition of the applicability domain, and the limitations of the test system regarding solvents used, solubility and interference in the test system, and ability to identify pro-haptens. Finally, the peer reviewers noted the low specificity of the EASA method as compared to human data and stressed the need to assess multiple key events to accurately predict sensitizers.

Comments from the peer reviewers were considered carefully during finalization of the report.

[1] https://aspe.hhs.gov/hhs-information-quality-peer-review

[2] https://www.federalregister.gov/documents/2005/01/14/05-769/final-information-quality-bulletin-for-peer-review