

Summary Minutes

Independent Scientific Peer Review Panel Meeting

Evaluation of *In Vitro* Estrogen Receptor Transcriptional Activation Test Method for Endocrine Disruptor Chemical Screening

William H. Natcher Conference Center

National Institutes of Health

Bethesda, MD

March 29 - 30, 2011

Peer Review Panel Members:

John G. Vandenberg, PhD (Peer Review Panel Chair)	Professor Emeritus, North Carolina State University, Department of Biology, Raleigh, NC
A. John Bailer, PhD	Distinguished Professor and Chair, Department of Statistics Miami University, Oxford, OH
Christopher Borgert, PhD	President & Principal Scientist, Applied Pharmacology and Toxicology, Inc., Gainesville, FL
Grantley Charles, PhD, DABT	Senior Scientist, Toxicology Drug Safety Evaluation, Allergan Inc. Irvine, CA
Daniel Desaulniers, PhD	Research Scientist, Health Canada, Hazard Identification Division, Ottawa, ON, Canada
J. Charles Eldridge, PhD	Professor of Physiology and Pharmacology, Wake Forest University Health Sciences, Winston-Salem, NC
Susan Jobling, PhD	Professorial Research Fellow, Institute for the Environment, Brunel University, Uxbridge, Middlesex, UK
William R. Kelce, PhD	Principal, Exponent®, Cary, NC
Hyung Kim, PhD	Associate Professor, College of Pharmacy, Pusan National University, Busan, Republic of Korea
Steven L. Levine, PhD	Science Fellow and Product Safety Manager, Monsanto Company, St. Louis, MO

Peer Review Panel Members:

Alberto Mantovani, DVM	Head of Food and Veterinary Toxicology Unit, Italian National Health Institute, (ISS - Istituto Superiore di Sanità, viale), Rome, Italy
Ellen Mihaich, PhD, DABT	President and Principal Scientist, Environmental and Regulatory Resources, LLC, Durham, NC
Hiroshi Ono, PhD	Scientific Advisor, Hatano Research Institute, Food and Drug Safety Center, Kanagawa, Japan
Sherry L. Ward, PhD, MBA	Consultant, BioTred Solutions, New Market, MD
Marc Weimer, PhD	Biostatistician, German Cancer Research Center, Department of Biostatistics, Heidelberg, Germany
James Wittliff, PhD, MD <i>hc</i> , FACB	Professor, Department of Biochemistry & Molecular Biology; Director, Hormone Receptor Laboratory; Director, Institute for Molecular Diversity & Drug Design, University of Louisville, Louisville, KY
James Yager, PhD	Professor and Director of Toxicology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

ICCVAM and ICCVAM Endocrine Disruptor Working Group Members:

Surender Ahir, PhD	OSHA, Washington, D.C.
M. Cecilia Aguila, DVM	FDA, CVM, Rockville, MD
Don Bergfelt, PhD	EPA, Washington, D.C.
Michael Bolger, Ph.D, DABT	FDA, CFSAN, College Park, MD
Jeffrey Bray, PhD	FDA, CDER, Silver Spring, MD
Paul Brown, PhD	FDA, CDER, Silver Spring, MD
Kent Carlson, PhD	CPSC, Bethesda, MD
Warren Casey, PhD, DABT	NIEHS, NTP, Research Triangle Park, NC
Karen Davis-Bruno, PhD	FDA, CDER, Silver Spring, MD
Kenneth Delclos, PhD	FDA, NCTR, Jefferson, AR
Charles Eirkson, PhD	FDA, CVM, Rockville, MD
Ross Filice, MS	FDA, CDER, Silver Spring, MD
Suzanne Fitzpatrick, PhD DABT	FDA, OC, Silver Spring, MD
John R. "Jack" Fowle, III, PhD, DABT	EPA, OPP, Washington, D.C.

Kevin Gaido, PhD	FDA, CVM, Rockville, MD
David Hattan, PhD (<i>Chair</i>)	FDA, CFSAN, College Park, MD
Jerrold Heindel, PhD	NIEHS, OSTB, Research Triangle Park, NC
Huixiano Hong, PhD	FDA, NCTR, Jefferson, AR
Abigail (Abby) Jacobs, PhD	FDA, CDER, Silver Spring, MD
Devaraya Jagannath, PhD	FDA, CVM, Rockville, MD
Hajime Kojima, PhD	JaCVAM, NIHS, Tokyo, Japan
Annette McCarthy, PhD	FDA, CVM, Rockville, MD
Leslie McKinney, PhD	FDA, CDER, Silver Spring, MD
Catherine Richter, PhD	DOI, CERC, Columbia, MO
Jesudoss Rowland	EPA, Washington, D.C.
William Stokes, DVM, DACLAM	NIEHS, NTP, Research Triangle Park, NC
Julius Thigpen, PhD	NIEHS, Research Triangle Park, NC
Donald Tillitt, PhD	DOI, CERC, Columbia, MO
Thomas Umbreit, PhD	FDA, CDRH, Silver Spring, MD
Jon Wilkes, PhD	FDA, NCTR, Jefferson, AR

Invited Experts:

Joseph Haseman, PhD	Statistical Consultant, Raleigh, NC
---------------------	-------------------------------------

**European Centre for the Validation of Alternative Methods, ICCVAM EDWG
Liaison:**

Susanne Bremer, PhD	ECVAM, EC-Joint Research Centre, Ispra, Italy
Elise Grignard, PhD	ECVAM, EC-Joint Research Centre, Ispra, Italy

Public Attendees:

Public Attendee	Day		Affiliation
	Attended 1	Attended 2	
Patience Browne		√	EPA, Washington, D.C.
Kent Carlson	√		CPSC, Bethesda, MD
George Clark	√	√	Xenobiotic Detection Systems, Durham, NC
Abigail Jacobs	√	√	FDA/CDER, Silver Spring, MD
Susan Laessig	√		EPA, Washington, D.C.
Raymond Lewis	√	√	Thermo Fisher Scientific, Pittsburgh, PA
Scott Lynn	√	√	EPA/OSCP, Washington, D.C.
Ruth McMillan	√		Xenobiotic Detection Systems, Durham, NC
Thomas Umbreit	√		FDA/CDRH, Silver Spring, MD
Kate Willett	√	√	People for the Ethical Treatment of Animals (PETA), Norfolk, VA
Cathy Yang	√	√	CertiChem, Inc., Austin, TX

NICEATM:

RADM William Stokes, DVM, DACLAM Director
 Warren Casey, PhD, DABT Deputy Director
 Debbie McCarley Special Assistant to the Director

Support Contract Staff— Integrated Laboratory Systems, Inc. (ILS):

David Allen, PhD Steven Morefield, MD
 Patricia Ceger, MS Michael Paris
 Jonathan Hamm, PhD Catherine Sprankle
 Linda Litchfield Linda Wilson

Abbreviations: CDER = Center for Drug Evaluation and Research; CDRH = Center for Devices and Radiological Health; CERC = Columbia Environmental Research Center; CFSAN = Center for Food Safety and Applied Nutrition; CPSC = U.S. Consumer Product Safety Commission; CVM = Center for Veterinary Medicine; DOI = U.S. Department of the Interior; EC = European Commission; ECVAM = European Centre for the Validation of Alternative Methods; EPA = U.S. Environmental Protection Agency; FDA = U.S. Food and Drug Administration; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; ILS = Integrated Laboratory Systems; JaCVAM = Japanese Center for the Validation of Alternative Methods; NICEATM = National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods; NIEHS = National Institute of Environmental Health Sciences; NIHS = National Institute of Health Sciences; NLM = National Library of Medicine; NTCR = National Center for Toxicological Research; NTP = National Toxicology Program; OC = Office of the Commissioner; OPP = Office of Pesticide Products; OSCP = Office of Science Coordination and Policy; OSHA = Occupational Safety and Health Administration; OSTB = Organ Systems Toxicology Branch; PETA = People for the Ethical Treatment of Animals.

TUESDAY, MARCH 29, 2011

Call to Order and Introductions

Dr. Vandenberg (Peer Review Panel Chair) called the meeting to order at 8:30 a.m. and introduced himself. He then asked all Peer Review Panel (Panel) members (those present and those in attendance via teleconference), Dr. Stokes, and Dr. Casey to introduce themselves and to state their name and affiliation for the record. Dr. Vandenberg stated that there would be ten public comment periods and he asked that those individuals interested in making a comment register at the registration desk. He requested that the public attendees hold questions and comments until the conclusion of the Panel's discussions. Dr. Vandenberg emphasized that the comments would be limited to seven minutes per individual per public comment session and requested that all comments should be brief and succinct. He deferred other introductions of those in attendance from the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), the ICCVAM Endocrine Disruptor Working Group (EDWG), Integrated Laboratory Systems, Inc (NICEATM staff), and members of the public until after the presentation of introductory remarks by Dr. Stokes.

Welcome and Opening Remarks from the NICEATM Director

Dr. Stokes of the National Institute of Environmental Health Sciences (NIEHS) and Director of NICEATM welcomed everyone to the Panel meeting being held on the main campus of the National Institutes of Health (NIH) in Bethesda, Maryland. He explained the NIH role as the nation's biomedical research agency and described the NIH budget and affiliated institutions. He stated that the NIH mission is to conduct science in the pursuit of fundamental knowledge about the nature and behavior of living systems, to apply that knowledge to extend healthy life, and to reduce the burdens of illness, injury, and disability.

Dr. Stokes stated that the NIH as well as the 15 agencies that are represented on ICCVAM greatly valued the expertise and input of the Panel on its review of the science behind these test methods. He said that ICCVAM is required by law to evaluate new methods and to make recommendations about their scientific validity, their usefulness, and limitations. The scientific peer review is incorporated as a critical and essential part of that evaluation process.

Dr. Stokes provided a brief overview of ICCVAM and NICEATM, and identified the 15 Federal agencies that comprise ICCVAM. He summarized the purpose and duties of ICCVAM (as described in the ICCVAM Authorization Act of 2000¹), noting that ICCVAM, as an interagency committee, does not carry out research and development or validation studies. Instead, ICCVAM, in conjunction with NICEATM, carries out critical scientific evaluations of the results of validation studies for proposed test methods to assess their usefulness and limitations for regulatory testing, and then makes formal recommendations to ICCVAM agencies. Dr. Stokes said that since the first recommendations were issued by ICCVAM in 1999, the committee and its members have reviewed and evaluated over 40 alternative safety-testing methods that have subsequently been endorsed by U.S. and/or international agencies. These methods, if they are used, can typically reduce, refine, and in some cases replace animal use for required regulatory testing. Most of these methods have been adopted as international test guidelines or incorporated into international guidance documents. NICEATM is a center of the National Toxicology Program (NTP) and is headquartered at NIEHS located in Research Triangle Park, North Carolina. It conducts and coordinates international validation studies and provides administrative and scientific support for ICCVAM. The NTP coordinates toxicology testing programs across the federal government.

¹ http://iccvam.niehs.nih.gov/docs/about_docs/PL106545.pdf

Dr. Stokes also stated that ICCVAM provides recommendations on how to advance new science and technology into standardized test methods by holding workshops. Experts are asked to provide advice on aspects that might improve or advance the methods that they review, so that the scientific community can move forward with methods that use few or no animals, can be conducted more efficiently, and can provide more predictive data.

Dr. Stokes defined basic validation as a determination of the usefulness and limitations of a test method for a specific purpose. Validation is more formally defined in ICCVAM documents as the process by which the reliability and relevance of a test method are established for a specific, defined purpose. Reliability is a measure of the extent to which a test method can be performed reproducibly within and among laboratories over time, and relevance is defined as the extent to which a test method will correctly predict or measure the biological effect of interest. Adequate validation of a new test method is a prerequisite for consideration for use in regulatory decision-making by federal agencies. The law specifically states that agencies must determine the method to be valid for its intended purpose prior to requiring, recommending, or encouraging its use.

Dr. Stokes explained that the last step in evaluating the validity of a test method is independent scientific peer review. The proposed use of the test method must provide for equivalent or improved protection of human and/or animal health or the environment compared to the method that it is proposed to replace or be used in place of. Once validation studies are completed, then draft documents called background review documents (BRDs), as well as draft test method recommendations, are prepared by ICCVAM (through NICEATM). These are made available to the public and provided to the peer review panel. During the independent peer review an independent report is generated and made available to the public and to ICCVAM's advisory committee for comment. Finally, ICCVAM considers the Panel report, the comments from the public, and comments from the advisory committee and then develops a final test method evaluation report that is transmitted to federal agencies and, where appropriate, forwarded for international consideration.

Dr. Stokes concluded his opening remarks and thanked the Panel members for their commitment of expertise, time, and effort and acknowledged their important role in the ICCVAM test method evaluation process.

Overview of the ICCVAM Evaluation

Dr. Stokes provided an overview of the ICCVAM evaluation process for the validation study. He told the Panel that they would review the validation status of an *in vitro* endocrine disruptor assay (BG1Luc ER TA Test Method, hereafter known as BG1 method) to detect whether chemicals could interfere and interact with the estrogen receptor. He provided the audience with a brief discussion of how endocrine disruptor substances could interfere with the normal function of endogenous hormone signals, which can lead to abnormalities that have been shown in laboratory studies in terms of growth, development, and reproduction.

Process and Charge to the Panel

Dr. Stokes explained to the Panel that the BRD provided a comprehensive compilation of all the information and validation data supporting the validity of the BG1 method. The duties of the Panel included review of the BRD for its adequacy and completeness and then consideration of the draft test method recommendations and the extent that this documentation supported those recommendations.

Dr. Stokes reviewed the charge to the Panel:

- Review the draft BRD for completeness and to identify any errors or omissions
- Evaluate the information in the draft BRD to determine the extent to which each of the applicable criteria for validation and acceptance have been appropriately addressed

- Consider the draft test method recommendations and comment on the extent to which they are supported by the information and data in the BRD. Those recommendations address the test method uses and limitations, recommended standardized protocols, recommended test method performance standards, and proposed future studies.

Dr. Stokes provided a brief timeline of the evaluation process. This process included publication of a public notice announcing a meeting and the availability of all the materials that have been provided to the Panel for review, implementation of Panel subcommittee meetings to determine initial draft positions, and reception of public comments for consideration. He explained that following the peer review meeting, the Panel would prepare and agree on a final report, and that Dr. Vandenberg would make a presentation at the June meeting of ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods (SACATM). ICCVAM and its Endocrine Disruptor Working Group (EDWG) will consider all this information, finalize its recommendations, and prepare the test method evaluation report for transmittal to federal agencies in the fall of 2011.

Dr. Stokes acknowledged the ICCVAM committee principal, alternate, and other representatives from the various agencies, as well as the participants on the EDWG from the various agencies. He also cited the international liaisons from ECVAM, the Japanese Center for the Validation of Alternative Methods (JaCVAM), and the NICEATM staff who organized the meeting under the leadership of Dr. Warren Casey.

Conflict of Interest Statement

Dr. Stokes reminded the Panel that they were meeting as an NIH special emphasis panel, under that charter, and he indicated that he would serve as the Designated Federal Official for the public meeting. He then read the conflict-of-interest statement and again asked members of the Panel to identify any potential conflicts for the record. Dr. Vandenberg asked the Panel members to declare any direct or indirect conflicts based on Dr. Stokes' statements and reminded the Panel that everyone had already signed a conflict-of-interest document.

Overview of Agenda

Dr. Vandenberg provided a general review of the agenda and outlined the process for reviewing each of the topics. Dr. Casey was to present the overview and background of the BG1 method and the validation study to the Panel on the first day of the meeting. Following that presentation, the Panel would break up into four separate groups with each group reviewing a specific area of the study. The leader of each group would then present the results of his or her group's discussion to the entire Panel when the whole Panel reconvened. All the panelists would have the opportunity to read the entire BRD, and those with specific areas of interest would have their comments discussed with the entire Panel. This would lead to a Panel discussion on the test methods, the validation of the data and the results, and the accuracy of the test method. The second day of the meeting would include Panel discussions on the test method reliability, the other studies that have come up in the interim, the animal welfare aspect, and other practical considerations. The Panel would also discuss the ICCVAM draft recommendations, the usefulness and limitations of the test methods, and potential future studies. Dr. Casey would present a summary of the BG1 method, and the Panel would discuss the test method performance standards. After each of these discussions and presentations, the public would have the opportunity to ask questions or make brief presentations.

Overview of the BG1Luc ER TA Test Method (LUMI-CELL® ER)

Historical Background

Dr. Casey provided the background for development of endocrine disruptor testing. He stated that the Federal Food, Drug, and Cosmetic Act sent a mandate to the U.S. Environmental Protection Agency (EPA) to develop a screening program, using appropriately validated methods or other scientifically

relevant information, to determine whether certain substances may have an effect in humans that are similar to an effect produced by naturally occurring estrogen. The Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) of the EPA recommended a two-tier screening program:

- Tier 1 – five *in vitro* and six *in vivo* tests, to identify substances with the potential to interact with the estrogen, androgen, or thyroid systems
- Tier 2 – a series of *in vivo* tests

In January 2004, Xenobiotic Detection Systems (XDS) nominated their LUMI-CELL[®] ER assay for a validation study to ICCVAM and in October of 2004, ICCVAM and SACATM considered the assay as a high priority method for validation. He described the prioritization of the *in vitro* assay based partly on the following factors:

- The method is faster and cheaper than any *in vivo* method
- A concentration-response curve is obtained from this method that is not always available in an *in vivo* study
- This method has an advantage over the currently existing *in vitro* method in that it has an agonist component and an antagonist component

Dr. Casey provided other prioritization criteria for *in vitro* test methods:

- They should be applicable to multiple agencies or programs
- They should be amenable to a high-throughput format as part of the Tox21 effort
- They should be applicable to the Organisation for Economic Co-operation and Development (OECD) conceptual framework, which is a weight-of-evidence approach to assess ED potential

He said that any assay that can be developed to help assess the endocrine disrupting potential of chemicals would add value to protecting human health. Dr. Casey noted that ICCVAM and NICEATM concluded that the test method is applicable to the criteria required by the agencies though the agencies are not obligated to accept.

Dr. Casey spoke on the potential of the BG1 method to reduce, refine, and replace animals. There is no direct replacement or refinement for the use of animals, but the BG1 method could be used as a substitute or an alternative test to the CERI STTA assay (Stably Transfected Human Estrogen Receptor- α Transcriptional Activation [STTA] assay validated by the Japanese company Chemicals Evaluation and Research Institute [CERI]). The BG1 method data showed that the assay has 100 percent specificity for detecting compounds that bind the estrogen receptors. Dr. Casey suggested that the BG1 method could possibly replace the direct uterine cytosol assay, which, although is an *in vitro* assay, does use animals as a source of ER. The BG1 method also has 100 percent sensitivity with the very small data set from the rat uterotrophic assay (i.e., any compound that tested positive in the uterotrophic assay was positive in BG1 method). Dr. Casey stated that the only validated ER TA method in use is the CERI STTA assay (an agonist only test), also known as OECD Test Guideline 455, which was directly adopted by the EPA for their EDSP program.

The BG1 method uses ovarian cancer cells that have both endogenously expressed ER-alpha and beta and cells that naturally have these receptors have the machinery in place to process signaling from those receptors. The BG1 method can identify antagonists, which creates the potential to identify a wider range of chemicals.

Validation Study Design

Dr. Casey said that a highly detailed standardized test method protocol was developed from October 2004 to October 2005 and ICCVAM recommended conducting an international multi-laboratory validation study. The study was organized by NICEATM, in conjunction with ECVAM and

JaCVAM, and one laboratory was identified to perform the study in association each of the three respective centers. A four-phase validation study was initiated. Phase 1 generated historical data and established that laboratories could adequately handle the reference standards and controls. Three replicate tests were run with four coded substances for agonism and four coded substances for antagonism during Phase 2a, the laboratory-qualification phase. Eight coded substances were tested for agonism and eight coded substances were tested for antagonism during Phase 2b, the qualification phase. Phase 3 included the testing of 41 coded substances one time each at all three laboratories. Only XDS participated in Phase 4 where 25 coded substances were tested. Additional retesting at XDS was completed in June 2010.

Background Review Document (BRD)

The BRD was drafted in 2010 (October through December) and reviewed by ICCVAM. The Federal Register (FR) notice for the peer review meeting was published in February 2011 and the BRD was made available to the public and the Panel.

Panel Discussion

Dr. Kelce initiated the Panel's discussion of Section 1 (Introduction) of the BRD for the BG1 method by noting that some typographical errors needed correction in the BRD and recommended that Phase 2 of the validation study should be known as the laboratory proficiency phase instead of the qualification phase. The Panel suggested including a more thorough discussion of the specific advantages of a transactivation assay relative to other *in vitro* ER assays such as providing advantages compared to binding assays since they examine estrogen agonist and antagonist activity, and detecting endocrine-active substances that potentially can act through non-receptor-mediated mechanisms. BG1 cells express both ER α and ER β and as such, possess the transcriptional machinery required for estrogen responsiveness.

BG1 cells have been stably transfected with four copies of the ERE and a luciferase reporter system. The Panel asked for clarification as to why there were four copies of the EREs and why the EREs are in their current orientation. Stably transfected cell lines have advantages over transiently transfected cell lines, including long-term utility without the need to transiently transfect cells each assay. The BG-1 test method has demonstrated responsiveness to estrogens and limited cross-reactivity with ligands of other steroid hormone receptors. Each of these important points should be emphasized in BRD.

Dr. Kelce asked whether ICCVAM's prioritization criteria and regulatory requirements were adequately discussed. The Panel was satisfied with the criteria discussion but agreed that the regulatory requirements need to be definitively detailed by their respective agencies worldwide. The specific regulatory use of data generated with this method has yet to be specifically defined (also applicable to the CERI STTA). Accordingly, it is essential that answers to the following questions be provided before making definitive conclusions regarding the usefulness and limitations of this assay:

- Is the BG1 method going to be added to the EPA Tier 1 battery?
- Will the BG1 method replace the STTA assay (if considered)?
- Will the BG1 method be used as a stand-alone screening assay for estrogen agonists and antagonists (i.e., replace the binding assay)?
- Will the BG1 method be developed into HTS screening assay?

The BRD should indicate that these issues are ill defined and that the agencies should provide input.

Dr. Borgert stated that ICCVAM criteria implied that all assays had to be validated for a specific purpose. However, the BG1 method has no defined use other than it is used in the same way as CERI STTA (which is also ill-defined as to the true use of the method). He said that one is trying to find substances that are ER agonists, and the definition of agonist is one that is positive in the CERI

method, then circular reasoning is being used. He asked that more detail be provided to adequately address the validity of the method.

Dr. Borgert expressed concern as to how these assays would be used. He stated that the assays could not be adequately compared until the specific use of the test method is defined. He proposed the following questions:

- If a substance is positive for ER agonism, does that mean it has the potential to be estrogenic *in vivo*?
- If the assay cannot provide information on whether a substance will be active *in vivo*, what is the meaning?
- Do we believe that androgens can act at the ER and is that relevant?

Dr. Borgert said that his comments were not criticizing the effort, but suggested that the Panel was compromised by the lack of specificity by the intended use of the methods.

Dr. Yager stated that the assay includes a qualitative response (positive or negative) and a quantitative response. The positive or negative does not consider concentration response even if it is near 1 mM or 100 μ M, which would make it more or less irrelevant *in vivo*. He said that determination of an EC value misses what concentration is available to the cells. Cell culture conditions will affect absorption of compounds to different degrees, depending upon their structure and chemical properties and there is no proposal to determine what the free concentration is that is available to the cells. He stated that a second element is agonist versus antagonist and noted that the BRD states that *in vitro* transcription systems have the potential to detect antagonists. This is confusing and is another element of specificity. He asks if the compound is actually acting as an estrogen or whether the second element of specificity is antiestrogen.

Dr. Casey clarified that the CERI STTA assay has never been validated to detect antagonists, though in theory, it should be able to. He said that interpretation of these assays is more than just positive and negative. EC₅₀ values are collected but comparison of positive or negative to EC₅₀ values was not part of the validation study design. This study showed how the BG1 method related to the other transcriptional assay and to the binding assay and how EC₅₀ values generated in this assay correspond to published EC₅₀ values. The Panel agreed that more commentary is needed on the potency and dose-response since IC₅₀ and EC₅₀ data are collected according to the protocol.

Dr. Levine stated that the test system can also detect endocrine-active substances through non-receptor-mediated mechanisms and the EPA examines this in their matrix to detect potential activity. He accepted that the BG1 method could potentially detect non-receptor-mediated mechanisms and considered what types of diagnostic tools are available for such situations and how one would address that in the context of the screening program that many compounds are going through. Dr. Levine suggested that this issue and the accuracy and specificity aspects be discussed in the BRD.

Dr. Borgert said that EPA Test Guideline 890.1300 states that the aim of the transcription-activation assay is to evaluate the ability of a chemical to function as an ER ligand and activate an agonist response for screening and prioritization purposes but can also provide mechanistic information that can be used in a weight-of-evidence approach. He also said that if an estrogen agonist is being defined as one that is positive in this assay, then it is very much a circular reasoning process. The assay is validated for estrogen agonists that are defined by the fact that they produce a positive in this assay or in any transactivation assay and then the assay is deemed good because it identified the compounds that it identified. The real problem is there is specificity on what this means or how it will be used. Dr. Borgert requested that the document should state this conundrum. Dr. Kelce added that a positive result does not determine whether the result is ER-mediated or not. The only way to tell is to do a co-incubation with ICI. Another positive result will indicate that it is not receptor-mediated and this is the only way to know whether that satisfies that specific criterion mentioned in the CERI

STTA assay. Dr. Borgert countered that if an estrogen agonist is defined as a compound that gives a positive response in a transcription-activation construct then he agrees with Dr. Kelce. However, he also said that if an estrogen agonist is actually a compound that is a ligand at estrogen receptors and produces the cellular and physiological responses of an estrogen, then the two might not match exactly. This is observable with some of the compounds that produce a positive result in these assays, albeit at a potency so far removed from the natural ligand that they could never actually function as an estrogen in an organism.

Dr. Stokes stated that the test methods were screening assays that provide mechanistic or pathway-type information and that these types of assays will be prevalent in the future in the field of toxicology. He said that obtaining many nonspecific positive responses that are not ligand-mediated in a screening assay makes the usefulness of that assay questionable. He requested that the Panel provide insights on the likelihood that mechanisms other than a ligand interaction caused that positive response and how often this might occur. Dr. Borgert stated that the literature shows how the wrong ligand activates the receptor and can see that there are substances that legitimately are activating the receptor, but at concentrations that are not at all relevant.

The Panel agreed that the purpose and mechanistic basis of the BG1 method was adequately described with the caveats mentioned previously regarding vector construct and design. However, the Panel stated that the use of the proposed test method in an overall strategy of hazard or safety assessment is unclear. The BRD should indicate that this has yet to be defined. Additionally, the Panel asked that relevant regulatory agencies respond with answers to the previous questions to more clearly define this issue and suggested that the BRD should propose how these data should be used for safety assessment. Regulatory agencies have yet to define how individual assays within the Tier 1 battery will impact safety assessment.

Dr. Stokes stated that the draft recommendations provided to the Panel include proposed recommended uses and limitations for a specific use. The Panel can include information in the meeting report on demonstrated other uses or potential other uses that need further data, either from this test system or other test systems, to characterize that usefulness. Dr. Casey added that it would be very useful for the Panel to provide an assessment of the test method (e.g., is this method as good as the current test, should you get the same results, can a company use this to submit data to the EPA, is there enough data to replace the receptor binding assay). Additionally, recommendations on additional potential applications or suggestions for other validation studies would be welcome.

Public Comments

Dr. George Clark (Xenobiotic Detection Systems [XDS]) explained why the BG1 cells were used for the validation study. He stated that XDS wanted to develop a naturally responsive estrogen receptor in a human ovarian carcinoma along with four or five other different receptors to have a cell available that mimicked normal cells. He also said that the BG1 method has six logs of responsiveness and that you can obtain dose-response data and relative potency between the different systems. Dr. Clark also expressed that the EC₅₀ values were an important part of the validation because they could be used for dose setting and thereby reduce animal usage.

Dr. Kate Willett (PETA) spoke about the OECD performance-based test guideline concept of which the CERI STTA assay and, hopefully, the BG1 method will form the basis. This is a concept to expedite the validation of methods that are considered similar, i.e., they measure the same endpoint and use similar technology. In the review of the BG1 validation study, there will be the need to harmonize the two studies in terms of their specificity and sensitivity and output.

Additional Panel Discussion

Dr. Casey stated that he was on the OECD workgroup that was reviewing the performance-based test guideline concept and was aware of Dr. Willett's information.

The BG1Luc ER TA Test Method BRD Protocol

Dr. Casey presented an overview of the agonist and antagonist protocols that were developed for the validation study. He explained that the protocols were fine-tuned to be used by highly trained laboratory personnel because of the extreme level of sensitivity the test method is capable of achieving. Implementation of the protocols lends themselves to high variability. The agonist assay is a gain of function, meaning that when the ligand binds to the receptor, the receptor then binds to estrogen response elements and turns on the gene with an endpoint as a gain in signal. The antagonist assay is a loss of function where one starts with the cells in an induced state, increasing concentrations of test substance are added, and determination of whether there is a decrease in signal is the endpoint. The inherent problem is that cells that die produce a positive response and this is the number-one concern with this assay.

Both protocols have essential test method components: solvent controls, DMSO, reference standards (17-beta-estradiol for the agonist and raloxifene HCl for the antagonist) and two weak positive controls, which are several orders of magnitude in IC_{50} s or EC_{50} s from the reference standards. The maximum test substance concentration, unless otherwise limited by solubility or cytotoxicity is 1 mM for the agonist and 10 μ M for the antagonist. Seven concentrations at log10 intervals are tested in the range finder tests. EC_{50} or IC_{50} values are calculated using the four-Hill-parameter equation. Test acceptance criteria for the agonist assay include a minimum of threefold induction for the E2 standard. No data points are used where the visual determination of cell viability is less than 80%. The CellTiter-Glo ATP method for determining cytotoxicity shows that if you have less than 20 percent reduction in viability, then there is no effect on signal. The EPA steroidogenesis assay and the ISO9000 cytotoxicity test for medical devices allow visual observation of cytotoxicity.

Panel Discussion

Dr. Kelce led the discussion of Section 2 (Test Method Protocol) of the BRD. The Panel discussed the importance of a reliable cytotoxicity test. Agreement on the definition of cytotoxicity is difficult to achieve since reduction in cell numbers may be attributed to many factors, e.g., mitochondrial dysfunction, oxidative stress, and redox imbalance. The measurement of luciferase activity is problematic since the determination of cytotoxicity is subjective and only highly experienced laboratory personnel can make accurate determinations. The Panel suggested that a cross validation of the luciferase method and another cytotoxicity test method should be conducted.

The Panel noted typographical errors in the draft BRD, which are detailed in Appendix A and that although improvements to the protocol(s) during the course of the study were explained in the BRD, a better assessment of their impact on study results is needed to improve the document. The Panel agreed that the protocols appear complete and adequate in detail for a laboratory to conduct the study. However, some of the details in the protocols suggest requirements that should be more generic (i.e., less specific).

The Panel stated that the critical aspects of the test method protocol are adequately justified and described in the BRD. The use of visual assessment of cytotoxicity is subject to operator inconsistencies and may not always accurately reflect the viability of cells in culture. An accurate, objective cytotoxicity method is most critical for antagonist assays. Additional efforts should be undertaken to validate the utility of this approach for future use. The Panel suggested that a wider set of substances with known mechanisms of cytotoxicity should be tested and quantitative cytotoxicity methods are needed for developing new *in vitro* ER assays. The Panel agreed that the reference standards and controls proposed for the agonist and antagonist protocols are appropriate, but that future studies should consider including confirmation assays using a pure estrogen receptor antagonist (e.g., ICI 182,780). This would confirm ER binding behavior.

Test acceptance criteria were discussed and the Panel agreed that the plate acceptance criteria defined in the agonist and antagonist protocols appeared adequate. However the criteria for acceptance based on the DMSO controls could be too stringent and perhaps better criteria could be defined. The Panel also agreed that the proposed decision criteria for identifying a qualitative positive or negative response in the agonist and antagonist protocols were appropriate. However, the Panel suggested that potency and intrinsic activity relative to a known endogenous ligand are critical components in determining whether a substance is truly positive (or not) and this is not addressed in the current decision criteria. The Panel recommended inclusion of this component to improve the utility of the assay. Other recommendations by the Panel included 1) recording EC/IC₅₀ values (or other values, e.g., EC/IC₂₀) and associating them with the qualitative “+/-” call, and 2) implementing a relative potency approach by comparing test substances to known reference substances.

Public Comments

Dr. Clark stated that one can visually inspect a plate for cytotoxicity in less than 30 seconds and that other cytotoxicity endpoint assays take 15 to 20 minutes. He said that the test method is not a rapid, fast assay providing good data when using endpoints other than visual observations. He expressed that the test method is a screening assay and that this aspect is reflected in the protocol.

Additional Panel Discussion

Dr. Levine responded that the term “screening” applies to the endocrine disruption screening program, but people have to recognize that many of the screens are really tests and speed is not the most essential aspect of the test method. He said that assays currently conducted in the ED program take nearly a month or longer to complete for the end phase and they include extensive biochemical, histological, behavioral, and reproductive measurements. He believes the same level of rigor is needed in the *in vitro* assays as the *in vivo* assays and each assay is as important as the other in terms of data quality, data accuracy, and data reliability.

Validation Study Reference Substances

Dr. Casey presented background information on the validation study reference substances. He stated that NICEATM and ICCVAM performed an objective and thorough retrospective analysis of the data and updated the database (established in 2002) in 2010. The analysis used literature to classify reference substances that were well documented in peer-reviewed journals that consistently reported the call for the substances as one way or the other. In 2002, ICCVAM recommended a list of 78 reference substances to use to assess four different methods: estrogen and androgen binding and TA. Not all of the substances were intended to be estrogen receptor reference chemicals. There are certain compounds that should be used for reference classification of estrogen-active compounds.

Dr. Casey said that the definition of accuracy used for this study is the degree of closeness between a test method result and an anticipated reference. The literature citations used provided human ER TA assay results. If the call was positive in those assays in the literature, then the call was positive for the BG1 method. The intent of this study was to show that this assay is as sensitive as other ER TA assays that have been previously published. NICEATM reviewed all papers identified in the literature searches and thoroughly evaluated the quality of data presented.

Two criteria were based on the literature review (a minimum of two studies): a substance was labeled as a positive if it was reported as positive for activity in more than 50 percent of the studies; a substance was labeled as a negative if it was reported negative in all studies, and a minimum of two studies. Other substances were labeled presumed positive or presumed negative as a convenience for classification. The only substances we used in the study were the positive and the negative. Based on those criteria, 42 substances were identified for use in agonist accuracy (33 positive, 9 negative) and 25 substances for use in antagonist accuracy (3 positive, 22 negative). Dr. Casey suggested that one of the clear concerns regards the three reference antagonists, which is a reflection of the literature, and

may be a reflection of the real or not-real danger of these compounds. He said that it is difficult to find consistent data for antagonists and that reports presenting compounds as being positive are false positives because the controls were not run correctly. Almost all compounds become positive after testing substances at 10 μ M in these types of assays.

One of the criteria for reference substances is that there should be chemical diversification. The reference substances used in this study fall in many different categories and are a good representation of different chemical structures. All of the reference substances had well-referenced activity one way or another.

Panel Discussion

Dr. Levine addressed the notion that this study was a retrospective validation. He said that this is not entirely different, in many respects, from the CERI STTA validation that was a pre-validation and retrospectively evaluated as a validation. Dr. Casey replied that Dr. Levine's understanding of validation may be a matter of semantics, and that this study was not designed as a retrospective validation study but as a prospective validation study. There was a retrospective analysis of the results.

The Panel noted typographical errors in the draft BRD, which are detailed in Appendix A and also cited relevant information that should be included in this section that would improve the document. Specifically, the Panel asked that the BRD reflect how ECVAM ensured that the in-house ECVAM laboratory did not know the identity of the controlled substances. The Panel agreed that a broad range of different chemical classes and physicochemical properties were tested as part of this validation exercise. The list follows ICCVAM guidance and is largely applicable to chemicals and products that are screened to evaluate the potential for estrogenic activity. Several classes were represented by as few as one substance. It is questionable whether these classes are sufficiently represented as no conclusions on usefulness or limitations specific to these classes can be made. The Panel recommended testing more substances in both the agonist and antagonist protocols. Future testing should include compounds from under-represented classes and compounds that have surface-active properties. Also, identification of additional known negative compounds for agonist activity is necessary since less than 25% of the agonist substances used for the accuracy analysis are negative. It is difficult to investigate false positives in a new test system if the majority of test chemicals are positive.

The Panel concluded that the use of the majority classification among results to establish the consensus reference classification assigned to each reference substance is a reasonable strategy. The Panel stated that a 50% cut-off for definitive classification is not very strong. If the quality of the data used was additionally considered then it would strengthen the rankings. However, for assessment of a screening assay, where perfect performance could not be expected or required, this method is sufficient. The Panel noted that the BRD should provide better clarity on how data quality in the literature review was evaluated/considered and it would be useful to include a Klimsch code or similar approach to evaluating published literature data quality. The Panel also recommended consideration of additional sensitivity analyses that use something other than the majority classification ranking.

Public Comments

Dr. Willett stated that the validation study evaluates the reliability or reproducibility of the assay and the use of compounds with limited information is equal to testing an unknown with an unknown. She noted that *in vivo* validation assays are not dependent on this range of reference compounds. The default assumption is that the animal study is applicable to all chemicals, unless there is a reason for it not to be and the *in vitro* method is the opposite. Perhaps there are physical properties of the assay

itself that may limit its applicability. The disparity between the two approaches is stark and may be prejudiced.

Validation Study Data and Results

Dr. Allen provided information on the validation study data and results. He stated that all records are publicly available and all of the original data were submitted to NICEATM as Microsoft Excel[®] and GraphPad Prism[®] files. They are readily available electronically for anyone. The study consisted of four phases, ultimately testing all 78 substances in at least one laboratory and testing 53 substances in three laboratories (52 in the Japanese laboratory due to Phenobarbital being unattainable). He discussed the DMSO control results and the E2 reference standard data. Fold induction was calculated on agonist test plates by dividing the highest average RLU value from the E2 reference standard by the average DMSO control RLU. Differences in solubility testing only impacted three substances in the data set, where androstenedione, 2-sec-butylphenol, and fluoranthene had some discordant results among the laboratories because of the starting concentration that was chosen for comprehensive testing. Cell viability assessment was not evaluated when assessing discordance among laboratories.

Dr. Allen stated that Phases 2 and 3 included agonist testing for 53 coded test substances at XDS and ECVAM (31 and 33 positive results, respectively; XDS had 10 negatives and 12 inadequates; ECVAM had 13 negatives 7 inadequates). Phase 4 included 25 additional substances to round out the list of 78 tested at XDS (7 positives, 14 negatives, and 4 inadequates). Antagonist testing resulted in very few positives, many negatives, two inadequates at XDS and ECVAM, and one at Hiyoshi. Among the 25 additional substances there are 4 positives, 20 negatives, and 1 inadequate.

Dr. Allen outlined the data quality aspects of the study. Both XDS and ECVAM conducted their studies according to Good Laboratory Practices (GLPs). The Japanese laboratory (Hiyoshi) used procedures based on principles of GLP that are outlined in the ISO 9000 standards. Laboratory reports contained the QA statements and any findings associated with those laboratory audits, and the NICEATM project coordinators also served as a secondary QA of data that were imported into the graphical software.

Plate acceptance criteria were established based on results generated in reference standards and control wells. Protocol standardization in the first phase of the validation study, as well as Phase 2, where historical data were generated, included acceptance criteria based on the historical databases. Dr. Allen said that high plate failures, 61 percent for agonists and 38 percent for antagonists, occurred in the first part of Phase 2 and some were an indicator of poor-quality data. The plate acceptance criteria were reconsidered, particularly the agonist E2 EC₅₀ and the methoxychlor RLU control value, in addition to the antagonist raloxifene IC₅₀ and the flavone control RLU values. The qualitative and quantitative outcomes for test plates that met all acceptance criteria versus those that failed to meet one or more of the acceptance criteria were considered. Analyses showed that there were test plates whether they met criteria or not, where the answer was the same, i.e., E2 EC₅₀ and the methoxychlor RLUs, in addition to, for the antagonist assay, the raloxifene IC₅₀ and the flavone control RLUs. These were dropped as acceptance criteria and the changes were incorporated in the subsequent phases of the validation study. The impact of changing those acceptance criteria showed a reduction in percent failures for Phase 2a (61 percent failures to fewer than 30 percent failures). Similarly, Phase 2b plate failure rates for the antagonist assay were reduced by half. Phases 3 and 4 showed marginally decreased failure rates.

Substances were classified as positive, negative, or inadequate. Inadequate data were identified based on those substances that did not meet decision criteria for either a positive or a negative response. This classification was always due to poor-quality data that could not be interpreted because of major qualitative and quantitative limitations. Normally these substances would be retested, but this exercise was actually retrospective.

Panel Discussion (Data and Results)

Dr. Levine suggested bridging the CERI STTA and BG1 methods under a common protocol and to verify the performance of the assay and the performance standards that are in the BRD, perhaps slightly changed based on recommendations of the panel.

The Panel noted typographical errors, suggested edits, and identified points needing clarification in Section 4 of the BRD, which are detailed in Appendix A. The Panel agreed that the data for studies used to evaluate the accuracy and reliability of the BG1 method appeared to be provided. The interlaboratory studies of the BG1 method conducted by NICEATM, JaCVAM, and ECVAM have been included in the accuracy and reliability assessments provided in the BRD. The Panel noted that the NICEATM-ICCVAM question asking if all known relevant data for all studies used to evaluate the accuracy and reliability of the BG1 method had been provided to the Panel is misleading. The Panel questioned if sufficient data from other screens that add to the support of a particular finding have been included, i.e., has there been enough consideration of results from ER binding, uterotrophic, etc. to address accuracy.

The Panel suggested that potency evaluation and relative comparisons of potency to reference substances would be useful and that suggested statistical tests of trend be included in the evaluation of a positive call. The Panel recommended that the test results should include: characterization of activity in addition to a positive/negative call, identification of the scale of measurement needed to compare tests, and evaluation of descriptive endpoints vs. continuous. Users of the test method should consider diagnostic testing as an example of comparing results from multiple tests. The Panel agreed that evaluation of the data resulting from the BG1 method requires further attention because analyses described often involve some transformation of the response relative to control responses and the variability in the control responses appears to be ignored in these constructions. The assumption of no downturn in the dose-response models implies that some preprocessing of the data points occurs to remove values that violate this pattern.

The Panel recommended that the criteria for an acceptable concentration response should be developed to only allow data that is sufficient for concentration response modeling. The current criteria of three points with non-overlapping standard deviations (SDs) are not sufficient. The Panel recommended additional discussion of the 4-parameter model to justify its use. The Panel stated that interpretation of the results should not rely solely on statistics but also on scientific judgment and should incorporate consideration of the nature and shape of the dose-response relationship and, if needed, the reproducibility of the response in independent experiments. Criteria should be established for acceptability of data to estimate potency values. The estimated values for EC₅₀ and IC₅₀ are presented as point estimates without any error being associated with them. There is uncertainty associated with these estimates, and it should be reported (e.g., confidence interval).

The Panel agreed that there was adequate documentation showing that coded substances were tested and experiments were conducted without knowledge of the identity of the substances being tested. This was critical for the evaluation of reliability and accuracy.

Public Comments

Dr. Willett stated that similar issues were discussed when interpreting the results from the CERI STTA validation study. The committee implemented a relative potency index that related the EC₅₀ of a test chemical to the EC₅₀ of 17-beta-estradiol. Instead of getting just a positive/negative result, there was a definite spread that could be divided into classes.

Panel Discussion (Data Quality)

The Panel noted typographical errors in Section 7 (Data Quality) of the BRD, which are detailed in Appendix A. The Panel suggested that the BRD should include: availability of audit results, statistical

evaluations, and methods for calculations; a time period for requiring this data be stored; recommendation of a secondary location for data backups.

The Panel agreed that the extent of adherence to national and international GLP guidelines for all submitted *in vitro* and *in vivo* test data and the use of coded substances and coded testing was adequately presented and that any deviations and alterations to original protocols including use of coded substances were relatively well described. However, formal training records of laboratory personnel were not available to confirm compliance with GLPs and Good Cell Culture Practices (GCCPs).

The Panel suggested that the BRD should include a description of the quality of the reference data, i.e., data from CERI STTA, data used to generate the ICCVAM reference classifications, data from the uterotrophic assay, and data from the rat ER binding assay. Consideration should be given to assigning greater weight to certain reference data, and a thorough description of the weighting methodology should be provided.

The Panel recommended that the responsible personnel for the quality assurance (QA) aspect of the study should be identified in the BRD. It was not clear to the Panel whether all appropriate QA steps have been performed. All audits should be documented in an appendix that is part of the BRD. The Panel asked that the BRD include additional explanation about the aspect of the updated classification system being developed after testing was complete and substances not being retested.

Test Method Accuracy

Dr. Casey provided background information on test method accuracy and the interpretation of results. A retrospective analysis was performed due to the way the decision criteria were originally laid out. The criteria were intended to estimate an effective lowest observed concentration based on a dose-response curve. The curve reaches a point that is statistically different than control, which equals a difference in response. The criterion for that threshold was three times the standard deviation of the DMSO control (three times above the DMSO baseline). Theoretically, the signal is above background. The characteristics of this type of classification system are that any value above that threshold constitutes a positive response. Any value that is statistically different than the DMSO control was a positive. If all the values were below that line, the test substance was a negative. There are some curves that cannot be classified because they are really close and those are called equivocal. Many of the curves should probably not have passed quality control, but the protocol did not allow the laboratories to exclude those. There was no quality metric around the data themselves. All the quality metrics were set around the DMSO controls, with the assumption that once the DMSO controls were adequate that should imply good-quality data for the rest of the plate.

Dr. Casey explained how the new decision criteria were determined. The two issues that were driving this issue were the high background and the variability with the data. Historical data were reviewed and the commonalities for clearly positive test results were evaluated. That evaluation showed that a general S-shaped curve with three non-overlapping data points on the slope was the common aspect. This was applied to all the other data, even the data that have high variability and if the substances were not positive by these criteria or negative by the other criteria, then were inadequate and the call could not be made. The criteria that were determined for what a positive should be in this assay is a dose response: positive slope, a peak or a plateau, and three data points with non-overlapping standard deviations. The other criterion, an amplitude of 2000 (20 percent of the E2 reference), is associated with the amount of noise at the baseline. Anything below that is negative response.

Panel Discussion

The Panel recommended that the primary comparison for the BG1 method should be the accepted reference method (CERI STTA) and this comparative analysis should be in Section 5.0 of the BRD. The Panel agreed that the ICCVAM reference consensus classification is an excellent additional

reference method for assessing the accuracy of the BG1 method and is essential due to deficiencies in the CERI data.

The Panel stated that there is no definitive way to assess the performance/accuracy of a screening assay. Therefore, this kind of novel comparison is useful, and the comparative results are good enough. The endorsement of a screening assay cannot be based on strict correlation to any one other method, but should be judged on a weight-of-evidence approach that includes all performance comparisons, the repeatability results, and scientific judgment regarding the biological relevance of the test system.

The Panel provided additional commentary on accuracy. BG1 method accuracy should be discussed in terms of how the BG1 method performance statistics compare to previously endorsed *in vitro* test methods. Demonstration of “agreement” between the two data sets (e.g., comparison of EC₅₀ and IC₅₀ values with reference data using regression) should use more appropriate methods than regression (e.g. Bland-Altman Plots and Limits-of-Agreement). Accuracy analyses should be based on results from individual labs along with a consensus classification determined based upon a “majority approach” using the three testing laboratories.

The Panel stated that a validation based upon unequivocal classifications might result in overly optimistic assessments of test method performance. The accuracy of the method in the validation should use unequivocal reference classifications from multiple test systems that have passed specific data quality measures, as was performed. This increases confidence in the results by eliminating questionable responses, positive or negative. The “accuracy” assessment for a screening assay can only be approximated, especially one without sufficient comparative data from a similar assay. Substances that result in equivocal reference classifications may provide additional insights in to aspects of the test method. Comparing the new test method with other methods is reasonable and indices of accuracy should be calculated only with data that meet certain quality control measures. If the data included in the ICCVAM classification do not meet these criteria they should not be used. Reference method data quality is an uncertainty for all of the reference methods used in the BRD because reference data quality is not provided in Section 7.0. Therefore, this criterion should not be imposed unfairly on these additional analyses. Use of data from other methods for comparison could be questioned since they may not be truly measuring the same biological impact. Accordingly, emphasis should be placed on describing the different purpose of the various tests, their advantages and disadvantages (the document should include a single section to provide these detailed explanations).

The Panel stated that assessment of biological impact is an admirable goal, but mechanisms are not fully understood. The indices of accuracy for assay performance have been used in all previous validation study analyses, and the BG1 method should not be subjected to different criteria. For example, the concordance of the BG1 method with the rat ER binding assay exceeds that for other assays endorsed as scientifically valid by validation authorities.

The potential for the assay to identify a greater portion of false positives (agonist testing) or false negatives (antagonist testing) than indicated in the results should be mentioned (due to low number of negative test substances). This is not seen as any reason for not endorsing the method for use in a test battery or other weight-of-evidence approach, but is important for users to understand. The contingency tables used to generate the summary statistics should be included in the document.

The Panel recognized that the original test criteria were inadequate and revisions had to be made throughout the study. Although this should not be a precedent going forward (optimal decision criteria should be selected *a priori*), the Panel acknowledged that changes made were performed in an appropriate manner.

The Panel agreed that the BG1 method is effective for generating data, in an amount and of a scientific value that is at least equivalent to the data generated from existing tests (i.e., U.S. EPA Endocrine Disruptor Screening Program Test Guideline OPPTS 890.1300: Estrogen Receptor Transcriptional Activation [CERI STTA]). The Panel concluded that the BG1 method protocol is an improvement over the CERI protocol given the extent to which the BG1 method protocol was standardized and optimized. Additionally, the Panel stated that ER binding assays indicate high concordance and therefore suggest that the outcomes of the stably transfected ER TA assay can provide reliable information about the biological effect of chemicals mediated by ER-ligand interactions. However, there appear to be sufficient data to consider replacing the rat uterine cytosol ER binding assay with the BG1 method. The Panel recommended that NICEATM and ICCVAM identify and use other available rat uterine cytosol ER binding data for these comparisons (i.e., substances beyond the 78 tested in the BG1 validation study). As an additional activity, the Panel recommended an evaluation of recombinant (human and animal) ER binding assays as a replacement for the rat uterine cytosol ER binding assay note.

Public Comments

Dr. Lynn congratulated the Panel for the expertise and diversity represented by the members. He stated that the rationale and discussion about the qualifications and the guidelines was well done.

Dr. Laessig addressed the Panel's discussion on whether or not to recommend replacing one assay with this one. This assay in particular is very oriented toward looking at human health because of the way the system has been designed. But the EPA screening battery also needs to take into account ecological effects. Some assays may be redundant but could provide useful information both human and ecological effects.

Additional Panel Discussion

The Panel continued general discussion of the various Panel-generated answers to the questions provided by NICEATM and ICCVAM concerning the information in the BRD. The Panel approved all of the answers subject to general wordsmithing.

Adjournment

Dr. Vandenberg adjourned the Panel for the day at 5:25 p.m., to reconvene at 8:30 a.m. on Wednesday, March 30, 2011.

WEDNESDAY, MARCH 30, 2011

Call to Order and Introductions

Dr. Vandenberg called the meeting to order at 8:30 a.m. and asked Panel members and other participants to introduce themselves. Dr. Stokes again read the conflict of interest statement and reminded the Panel that each person had signed the appropriate form.

Test Method Reliability

Dr. Allen provided a synopsis of the reliability section (Section 6) of the BRD. He defined reliability as a measure of the degree to which a test method can be performed and is calculated by intra- and inter-laboratory reproducibility and inter-laboratory repeatability. Reliability is based on the reproducibility of the reference standards and controls of the reference substances themselves within and between laboratories.

Dr. Allen discussed the intra-laboratory reproducibility of reference standards and controls, i.e., the agonist DMSO control RLU values and the antagonist DMSO control and E2 control RLU reference values. He explained that 12 substances were tested three times at all the three labs in Phase 2 and that there is 100 percent agreement within each laboratory for each of the three repeat tests for both agonists and antagonists. Similarly for inter-laboratory reproducibility, there was data from Phases 2 and 3 for substances that were tested in all three laboratories. Only those substances that produced a definitive result in at least two of the three laboratories were used for the reproducibility analyses. In agonist testing, two-thirds of the laboratories got the same answer with regard to positive-negative calls. There was 100 percent agreement across all labs for antagonist testing. For Phase 3, 83 percent of the substances had agreement across the laboratories in the agonist phase. For antagonist testing, most of the substances tested produced agreement among the laboratories, with only three instances showing discordance.

Panel Discussion

Dr. Mihaich initiated the discussion on test method reliability (Section 6 of the BRD) and presented the Panel with the questions proposed by NICEATM. The Panel concluded that not enough data were presented to ensure that a thorough analysis of intra- and inter-laboratory reproducibility had been adequately evaluated. Many of the figures in the BRD do not adequately present the variability associated with the test method and in fact may confuse variability with outliers and these sources of variability should be explored using appropriate methods, e.g., comparison of CVs and formal statistical analyses for evaluating variance components. The Panel stated that reproducibility had been addressed on three data levels: raw data (e.g., DMSO control measurements), derived endpoints (e.g., EC₅₀), and prediction of estrogenic activity. The analysis is descriptive but no statistical testing was performed which avoided sample size, power issues, and definition of equivalence margins. The Panel agreed that while this approach is acceptable, there are some issues that require additional discussion:

- Criteria used for determining what is an appropriate way to characterize intra- and/or inter-laboratory variability could include comparisons to established test methods that have been considered acceptable.
- Variability of EC₅₀ estimation needs to be summarized in Section 6 of the BRD. Summarizing EC₅₀ values from different compounds using plots, summary statistics, or agreement measures would provide further insight into the overall variability and reproducibility of the assay, e.g., compare the BG1 method reproducibility to reproducibility of similar test methods (e.g., CERI STTA).
- CERI STTA analysis of variance components provided components that could be useful for a direct comparison. This includes calculating the mean and standard deviation of

logEC₅₀ to complement the current analysis and to allow for a more direct comparison to the analyses of other assays that also included these calculations (e.g. CERI STTA).

- Sensitivity of analyses to assumptions that are made should be investigated.
- Assessment of reproducibility might be too optimistic by excluding substances from reproducibility analysis for which definite results were not determined.
- Rationale for selection of the substances used to evaluate intra- and inter-laboratory reproducibility should be more clearly described in the BRD.

Public Comments

Dr. Clark stated that there are variants with unknowns that you are trying to assay, and it is unknown exactly how the biological mechanism works. Members of the Panel discussed further variants such as DMSO controls, reference standards, and EC₅₀ values.

Additional Panel Discussion

Dr. Stokes added that ICCVAM has not published strict criteria that define what is acceptable reproducibility between and within a laboratory. There is inherent variability in test systems and variability due to differences in laboratory operations. The goal is to determine whether the variability is reduced sufficiently to get the same type of response back even with that background variation. ICCVAM does not define what is acceptable and not acceptable but wants to know how variability impacts the outcome of the assay.

Other Studies

Dr. Casey stated that Section 8 of the BRD provided summaries of other publications that related to the BG1 method. He said that numerous chemicals have been tested in this cell system, not necessarily with this particular protocol, in addition to the 78 chemicals that have been tested in this study. Although there are no published reports that this assay is being considered for validation, the test method is being evaluated in the Tox 21 effort (the high throughput testing 1536-well plate format) at NIEHS. Researchers are obtaining a threefold induction, and feel that the test method is highly transferable from 96-well plate format to the high-density plate format.

Panel Discussion

Dr. Borgert initiated the discussion on the BRD's presentation of other studies and presented the conclusions of the Panel. The Panel unanimously agreed that all the relevant data identified in published studies that employ this test method have been adequately considered. The Panel suggested obtaining QSAR-based predictions for ER binding from the literature on some of the validation study chemicals to evaluate a comparison to the BG1 and rat ER binding results.

Public Comments

There were no public comments.

Animal Welfare

Panel Discussion

Dr. Borgert provided the leadership on the discussion of animal welfare. The Panel concluded that further discussion was required to determine the extent to which the test method will reduce, refine, or replace animal use. In order to fully understand how this method will impact the 3Rs, there needs to be a better overview of the EPA EDSP Tier I screening battery and the proposed context into which this test method will fit in terms of the overall testing scheme. Additional discussion of the following topics would provide additional ways that a validated BG1 method could contribute to reducing animal use:

- Determine whether the BG1 method will be performed prior to- or simultaneously with- the other *in vitro* and *in vivo* assays
- Define the implications for possible Tier II testing, which will likely incorporate significant animal use

The Panel stated that comments provided in the EPA's 2011 budget state their intent to phase out the EDSP and to include high-throughput ED-detecting assays in the ToxCast screening battery, which should lead to reductions in animal use for regulatory testing in the long term. The BG1 method will probably contribute to the development of this screening battery. The Panel also determined that a major problem with *in vitro* ED testing and the EDSP screening battery is the lack of an *in vitro* method to evaluate metabolism. The inclusion of a metabolism component as part of any *in vitro/in silico* ED test battery will provide a more biologically relevant assessment of ED activity. Since there are many *in vitro* metabolism methods being used in drug development and being used for *in vitro* testing of other toxicity endpoints (e.g., genotoxicity), the importance of including metabolism as part of all future studies on *in vitro* ED assays should be mentioned in Section 9.0 of the BRD.

The Panel was concerned that implications of BG1 method validation in the EU and Japan were not mentioned in the BRD. ED chemicals are substances of very high concern (SVHC) in the REACH testing program, and validated *in vitro* methods have the potential to reduce animal use in these kinds of large testing programs. REACH-type programs are also being adopted by Asian countries, so the availability of validated *in vitro* and *in silico* methods to screen/prioritize chemicals for these testing programs has the potential to reduce animal use. Additionally, following validation, the development of *in vitro* ED assays into an OECD TG will broaden their potential for reducing animal use. The Panel agreed that concordance of the BG1 method with the rat cytosol ER binding assay of 97% and based on 33/34 test substances suggests the BG1 method can "replace" the rat ER binding assay within the EDSP Tier 1 battery (and similar test batteries). The Panel believes that the excellent concordance of the BG1 method exceeds the "performance" of other methods that have been endorsed as scientifically valid. Additionally, an assessment based on 34 test substances could provide sufficient confidence (power analysis could be conducted to confirm).

The Panel stated that the excellent concordance of the BG1 method with the rat uterotrophic assay (92%, 12/13) indicates that the BG1 method is an excellent candidate assay for replacing the uterotrophic assay, and thereby reducing animal use. The small data set, however, is not sufficient to recommend endorsement of the BG1 method as a replacement at this time. Therefore, the Panel could recommend the BG1 method as a high priority for additional studies. Retrospective analyses may be sufficient. If necessary, a prospective study could be conducted to further compare these methods, preferably by identifying additional materials already evaluated in the uterotrophic assay. *In vitro* metabolism must be included as part of the prospective component of this study. Also, since the BG1 method was already subjected to an extensive interlaboratory study, consideration for an abbreviated assessment (e.g., 1-2 laboratories) should be considered to reduce time and costs.

The Panel provided the following text as a possible concluding paragraph for Section 9 of the BRD:

The development of a battery of *in vitro* and *in silico* methods that can totally replace animal testing for detecting chemicals that have the potential to cause an adverse effect interact with the endocrine system is a biologically complex problem. For example, a method for the assessment of metabolized test substances need to be included with the *in vitro* assays, and assays for assessing the many modes of action of EDs on various tissues and species need to be developed and validated. The experience derived from validating and using the *in vitro* BG1 method is expected to contribute to our knowledge and promote progress toward this goal. It should lead to the broader use of cell-based methods for ED screening, and could include the use of cells from other species.

Public Comments

There were no public comments.

Practical Considerations

Dr. Casey stated that some of the practical considerations of the test method are 1) it should be considered as part of a weight of evidence approach even though it does not replace animal use, 2) it is a possibility that the test method may replace the rat uterine cytosol assay, and 3) it is extremely important to have well trained personnel that operate under a GLP environment.

Panel Discussion

Dr. Borgert led the Panel discussion on practical considerations and provided a list of practical considerations that should be addressed:

- Provide better clarification regarding the availability of the BG-1 cell line which has not been placed in a public repository; efforts should be undertaken to do so or to otherwise ensure the continued availability of the cell line.
- Employ a less subjective cytotoxicity assessment than the current visual method to improve transferability and implementation of the method by laboratories.
- Associate costs of equipment and supplies with the date they were acquired.

The Panel agreed that the level of training and expertise required to conduct the test method are reasonable for its wide use and that the protocol should recommend the need for adherence to GCCPs and GLPs in order to minimize variability.

Public Comments

Dr. Clark stated that one of the reasons that XDS developed the test method was to get dose response data that could be used to set doses in animal studies and reduce the number of animals used in range finding tests.

Dr. Willett stated that it might be appropriate for the Panel to make recommendations for improving the use of this BG1 method regardless of how it compares to the rat uterotrophic, assay, for example. One of those recommendations might be to explore the use of different metabolism systems with the BG1 method. This has implications more broadly for *in vitro* assays and might be included in recommendations for future studies. Additionally, considerations for measuring the actual concentration could be part of the recommendations. She also suggested that general recommendations that the Panel could make regarding changes in future validation exercises that might facilitate the validation of these types of assays. She acknowledged this kind of validation exercise cannot be physically be done for all of approximate 300 ToxCast assays but any recommendations in terms of improving the evaluation process would be very helpful.

Additional Panel Discussion

Panel members discussed the use of the acronym EDs (i.e., endocrine disruptors) and asked whether a new acronym should be used (e.g., EAC, endocrine active chemical).

Dr. Jacobs said that the use of the term ED and endocrine disruption for assays, which only assess possible interaction with the system, has caused the FDA problems. They should be called potentially interactive substances. She does not believe that endocrine disruptors and an adverse *in vivo* effect can be defined in an *in vitro* assay.

The BG1Luc ER TA Test Method Draft Recommendations on Usefulness and Limitations

Dr. Casey presented a restatement of what was contained in the BRD and listed the ICCVAM recommendations.

- ICCVAM proposes that the BG1 ER TA test method can be used as a screening test to identify substances with *in vitro* agonist activity. This use is based on an evaluation of available validation data and corresponding accuracy and reliability. ICCVAM concludes that the accuracy of this assay is at least equivalent to the current ER TA included in the regulatory testing guidance.
- ICCVAM proposes that the BG1 ER TA test method can be used as a screening test to identify substances with *in vitro* antagonist activity. This use is based on an evaluation of available validation data and corresponding accuracy and reliability.
- ICCVAM recommends that the protocols in the BRD that the Panel has reviewed should be used in any further testing with this cell system.
- There are some of the limitations that have been identified: test substances must be soluble in DMSO, cannot react adversely with DMSO or cell culture media, and cannot have endogenous luminescence or naturally inhibit luciferase activity.
- ICCVAM states that the concordance of this assay is similar with that of the rat uterine cytosol assay and the BG1 ER TA test method has potential replace the rat uterine cytosol assay.

Panel Discussion

The Panel agreed that based on the overall test method performance, as presented in the BRD, the BG1 method is capable of identifying substances as potential *in vitro* ER agonists and antagonists, provided the acceptance criteria described in the recommended test method protocol are met. In addition, accessibility to the cell line still needs to be resolved, which clearly can impact the future use of this assay, including as a screening test in a contract laboratory setting.

The Panel suggested that additional analyses could be performed that would strengthen the understanding of how well this screen performs compared to the current CERI STTA assay and within the battery as a whole. It is important that all efforts have been made to make use of the data that are available before additional laboratory work is conducted. The Panel agreed that the assay can be used as a screening test to identify substances with *in vitro* estrogen agonist properties and can be used in place of CERI STTA for regulatory testing but recommended additional reliability analyses (which may lead to revising draft ICCVAM recommendations pending their outcome).

The Panel also agreed that the assay could be used as a screening test to identify substances with *in vitro* estrogen antagonist activity although error rates may not be precisely estimated in the antagonist assay nor necessarily representative of the population of chemicals that may be tested.

Public Comments

There were no public comments.

The BG1Luc ER TA Test Method Draft Recommendations on Future Studies

Panel Discussion

Dr. Casey presented the Panel's recommendations for future studies. He stated that the Panel agreed that the available data supported the ICCVAM draft recommendations for the BG1 method in terms of the proposed future studies. In addition to the proposed studies, the Panel recommended that additional future studies should include:

- Efforts to validate the utility of the current cytotoxicity evaluation
- Searches for fully-defined media to replace FBS
- Attempts to account for compound metabolism/activation
- Conversion of the BG1 method (+/- metabolism) into a HTS assay format for use in ToxCast and other screening programs

The Panel stated that the concordance of the BG1 method with the rat cytosol ER binding assay suggests that the BG1 method and binding assays produce similar results. Regulatory agency clarification is needed to determine if both assays are necessary.

Public Comments

There were no public comments.

Additional Panel Discussion

Additional discussion included comments on who would actually perform future studies. Dr. Stokes stated that any recommendations relevant to advancing the usefulness and limitations of this assay are appreciated. Who carries out that work depends on who has the resources, as well as if it is a federal agency, where they have the statutory authority to do such work. Research agencies in the federal government can carry on R&D type work, but that does not preclude a non-government organization from performing such work as well. He said that the BG1 cell line is currently available through a commercial entity. Performance standards were developed for evaluating commercially available methods or those with some intellectual property rights associated with them. For agencies to endorse a method that is copyrighted, trademarked, or sold by a commercial firm, the basis by which that is considered acceptable must be stated so that anyone else can create a similar model and know what criteria it has to meet to also be considered acceptable. It is important to understand that perspective on the availability of the test. This method has been published in the open literature as to how the cell line was created, so that others could do that if they wanted to.

The BG1Luc ER TA Test Method Performance Standards

Dr. Casey stated that the performance standards are used to evaluate the accuracy and reliability of other proposed test methods, sometimes referred to as me-too methods. If they can meet these criteria, then they are considered functionally equivalent to the BG1 assay. The performance standards include the essential test method components, all critical aspects of the assay, a list of reference substances to be tested, and a defined set of accuracy statistics or of metrics to use for accuracy and reliability to insure that another method is the same as the current method.

The essential test method components include: a human ovarian cell line, stably transfected with luciferase; an appropriate solvent; test substance concentration up to one millimolar for agonist testing and up to ten micromolar for antagonist testing; a minimum of seven concentrations at log ten intervals; an evaluation of cytotoxicity; the use of reference standards, weak and positive controls, and solvent control.

He said that the interpretation of results is not a statistical approach but an empirical approach. If a positive curve is observed, then the EC₅₀ should be calculated if possible. The criteria for the antagonist tests include testing a ten-micromolar limit concentration and performing an IC₅₀ calculation if possible.

These are 34 proposed reference substances and all should be tested as coded materials. Performance of the BG1 method requires significant training to become proficient. In order to demonstrate equivalence, test as many chemicals as possible. All laboratory personnel should treat every chemical like it is E2 to avoid contamination issues and high variability. Discordant results with a new test compared to the BG1 method should be discussed in terms of the ability of the test methods to detect a similar range.

Panel Discussion (Performance Standards)

The Panel assessed the adequacy of the performance standards adequate for assessing the accuracy and reliability of test method protocols that are based on similar scientific principles and whether those methods measure the same biological effect as the BG1 method. The following concepts were proposed by the Panel and should be addressed in Section 1 of the BRD.

- The ‘intended purpose’ for the assay in the context of the screening battery has not been directly defined. The lack of clear purpose for the assay has profound implications for the setting of performance standards. Without a clear purpose, only provisional performance standards can be set. The detailed technical comments of the Panel should be understood in the context of this provisional nature. In turn, issues related to cytotoxicity concerns and assay performance, including the number of known negatives evaluated, are intimately related to the setting of performance standards and must be understood in the context of the provisional nature of the validation effort.
- A statistically significant difference between control and treated is an inadequate delimiter of agonist response. This is readily apparent from the chemicals deemed to produce positive responses. For example, both the 890.1300 assay and the BG1 assay are said to produce a positive response with methyltestosterone. Does this mean that methyltestosterone is to be considered an estrogen agonist, a ‘potential’ estrogen agonist, or something else? Or, is some distinction being made between “. . . the ability of a chemical to function as an ER α ligand and activate an agonist response, . . .” and “. . . the ability of a chemical to function as an ER α ligand and estrogen agonist.” In other words, has the validation effort made a distinction between a chemical ‘activating an agonist response’ versus actually functioning as an agonist? If so, what is the qualitative and quantitative relationship between a chemical that activates an agonist response and one that is actually an agonist? These questions need clarification before any meaningful validation can ensue.
- For example, the 20% cutoff for cytotoxicity might be a good standard if the purpose of the assay were to identify chemicals with potential to activate an agonist response, and an agonist response in this assay had been defined as intrinsic activity and potency at the ER α at least 25% and 5%, respectively, relative to 17 β -estradiol (percentages not to be taken as a recommendations, but merely for the sake of interjecting sufficient detail to make the point clear). With that type of clear, detailed definition of “agonist response,” it would then be possible to determine whether the performance criteria for the assay were adequate for assessing the accuracy and reliability of the test method.
- Intrinsic activity is the ability of a drug-receptor complex to produce a maximum functional response. Intrinsic activity is sometimes used interchangeably with efficacy; however, intrinsic activity refers to a cellular response whereas efficacy is more often used in the context of a clinical response. Assuming equivalent pharmacokinetic parameters and affinity, a drug with greater intrinsic activity would be expected to have greater efficacy. Affinity is a measure of how tightly a drug binds to a particular receptor, and is often defined by the dissociation constant. Potency is the intensity of effect produced per unit of drug, and is a function of intrinsic activity and affinity.
- Neither the BG1 nor 890.1300 assay measures a complete agonist response, defined as a chemical that binds to an estrogen receptor in a cell and triggers an estrogenic response by that cell. Instead, these assays measure the first two steps of an agonist response at the alpha subtype of the estrogen receptor (ER α) via an artificial construct that couples binding and activation of the receptor complex to a reporter construct, for the BG1 assay, the enzyme luciferase. Activation of luciferase is not a normal physiological response of estrogens; hence, this step is not considered a component of estrogen agonist activity.

Because neither the BG1 or the 890.1300 assay measures a bona fide agonist action, validation would require defining both qualitatively and quantitatively the relationship between the stated purpose of the method and the activity actually measured by the assay. This relationship has not been characterized, quantitatively or qualitatively, for either the BG1 or the 890.1300 series assay.

Panel Discussion (Essential Test Method Components)

The Panel determined that the justification for some of the essential test method components is not clear and recommended clarification of the following points be addressed in the BRD:

- The most critical point for the type of cell line used is that it should include the appropriate “machinery.” The specific tissue source, cell type, and species may not be critical.
- Explain why the maximal concentrations were suggested.
- Explain why log₁₀ spacing of seven concentrations is needed.
- Address the evaluation of cytotoxicity and how modeling of the possibility of such a response could also be considered.
- Explain why 20% of maximal response is the cut-point for significant response.
- Address how a test of negative trend might be preferred to the suggestion of “three points with non-overlapping error bars.”
- Define what the error bars correspond to (e.g., standard deviations, standard errors, half-width of confidence intervals). Use of non-overlapping confidence intervals is a conservative way to declare difference between parameter and better guidance is needed on this.

Panel Discussion (Reference Substances)

The Panel stated that in general, the criteria used to select the performance standards reference substances are adequate. However, one could question the appropriateness of defining reference substances as positive based upon >50% of ER TA studies indicating a positive response. Given that the quality of reference data has not been detailed the Panel cannot definitively determine appropriateness of the reference substances.

The Panel agreed that the list of reference substances upon which to evaluate the performance of functionally and mechanistically similar test methods would be considered adequate if there were more negatives and proportionally fewer positives in the list for agonist testing. The list has reasonable overall diversity and reflects the extensive effort to obtain relevant information. There are enough substances to lend sufficient robustness to an assessment. There is a good range of estrogenic activity over several orders of magnitude, as well as a few confounders to assess the robustness of the assay and methodology. There may be an opportunity to revisit the list of reference substances and make modifications based on experience gained in the assay. Including discordant chemicals on the reference substance list is important because they are critical for truly characterizing the limitations of the assay. The potent estrogens on the reference list should not be missed and there could be some tolerance for discordance for the weaker acting reference substances. Any discordant results should be discussed in terms of the ability of the test method to detect a similar range of potencies and intrinsic activity and chemical/product classes.

Panel Discussion (General Comments)

The Panel determined that the number of repeat experiments to evaluate intra- and interlaboratory reproducibility appeared to be a good starting point for evaluation of the test method. The intralaboratory assessment (based on at least three tests with 12 chemicals) and the interlaboratory reproducibility assessment through one trial (three laboratories) allowed for ‘real world’ evaluation of

the methodology on the validation reference set of materials. The Panel recommended that the BG1 method should use a range of accuracy (or perhaps the lowest %) and include a metric of potency and intrinsic activity in the accuracy evaluation (a measure of uncertainty needs to be included). There is also a need to establish tolerance to the vehicle (ideally at least 1%). The test report should include potency and sensitivity analyses based on EC₅₀ values (agonist) or IC₅₀ values (antagonist).

Adjournment

After the discussion, Dr. Hayes adjourned the Panel for the day at 5:31 p.m.