Contains Nonbinding Recommendations

Draft – Not for Implementation

Radiation Biodosimetry Devices

Draft Guidance for Industry and Food and Drug Administration Staff

DRAFT GUIDANCE

This draft guidance document is being distributed for comment purposes only.

Document issued on December 30, 2014.

You should submit comments and suggestions regarding this draft document within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. Identify all comments with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions about this document, contact the Division of Molecular Genetics and Pathology at 301-796-6179 or Jennifer Dickey at 301-796-5028 or via email at Jennifer.Dickey@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health

Office of *In Vitro* Diagnostics and Radiological Health
Division of Molecular Genetics and Pathology Devices

Preface

Additional Copies

Additional copies are available from the Internet. You may also send an e-mail request to CDRH-Guidance@fda.hhs.gov to receive a copy of the guidance. Please use the document number 1400045 to identify the guidance you are requesting.

Table of Contents

Radiation Biodosimetry Devices

Draft Guidance for Industry and Food and Drug Administration Staff

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. Introduction

FDA has developed this draft guidance to facilitate study designs to establish the analytical and clinical performance characteristics of radiation biodosimetry medical countermeasure devices. Radiation biodosimetry countermeasure devices are devices used for the purpose of reconstructing the ionizing radiation dose received by individuals or populations using physiological, chemical or biological markers of exposure found in humans. Radiation biodosimetry technologies may be used at various stages during triage and treatment after the exposure of a population to ionizing radiation as a result of intentional harm or as an unintended consequence of a disaster. Devices may be designed to give quantitative outputs or qualitative information around a clinical decision making cut-point. Likewise, devices may be designed for use in field triage settings, at patient bedsides, or in Clinical Laboratory Improvement Amendments of 1988 (CLIA) certified clinical laboratories. FDA considered both high-throughput and single-use devices in developing this draft guidance document.

This draft guidance document does not provide specific study designs; it describes design principles for studies that may be used to establish a reasonable assurance of the safety and effectiveness of radiation biodosimetry devices. Sponsors should develop a validation plan to establish the analytical, pre-clinical, and clinical performance characteristics in order to substantiate the claims in the device intended use statement, and discuss this plan with the FDA prior to beginning studies.

Throughout this guidance document, the terms “we,” “us” and “our” refer to FDA staff from CDRH. “You” and “your” refers to the applicant or sponsor.

FDA's guidance documents, including this draft guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a

topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidance means that something is suggested or recommended, but not required.

II. Background

A radiological event, for instance from use of an improvised nuclear device or radiological dispersal device or from a natural disaster, could potentially expose thousands of individuals to high levels of radiation that would require immediate assessment and medical intervention. A coordinated medical response including triage and treatment would be important to mitigate harm resulting from the unintended exposure of individuals to ionizing radiation. Radiation biodosimetry tools would be a critical component of such a response. Biodosimetry is a surrogate for knowledge of the absolute dose delivered to an individual. It allows for the assessment of the likelihood of a patient developing acute radiation syndrome (ARS) and to develop an appropriate treatment plan for the patient.

Most radiation biodosimetry devices are *in vitro* diagnostic devices (IVDs), as defined in 21 CFR 809.3(a).[1] However, a wide array of technologies may be employed to assess biological responses to radiation. Methodologies amenable to radiation biodosimetry devices could include nucleic acid based devices that utilize technologies such as polymerase chain reaction (PCR) or microarrays, devices designed to detect changes in protein expression using technology such as enzyme-linked immunosorbent assays (ELISA) or flow cytometry, and devices designed to detect other biological signals induced by exposure to radiation.

Sponsors who intend to market radiation biodosimetry devices must, in addition to other applicable requirements, conform to the general controls of the Federal Food, Drug, and Cosmetic Act (FD&C Act), and obtain premarket clearance or approval prior to marketing their devices. When finalized, this draft guidance document will represent our current thinking regarding the recommended design of studies to demonstrate a reasonable assurance of the safety and effectiveness of radiation biodosimetry devices for unintended exposures to radiation. We consider these recommended studies to be relevant for premarket notifications (e.g., 510(k) submissions or premarket approval applications (PMAs) that may be required for a particular device). General information about 510(k) submissions and PMAs are outlined in the following guidances:

- "Format for Traditional and Abbreviated 510(k)s" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm084365.htm)

[1] *In vitro diagnostic products* are those reagents, instruments, and systems intended for use in the diagnosis of disease or other conditions, including a determination of the state of health, in order to cure, mitigate, treat, or prevent disease or its sequelae. Such products are intended for use in the collection, preparation, and examination of specimens taken from the human body. These products are devices as defined in section 201(h) of the Federal Food, Drug, and Cosmetic Act (the act), and may also be biological products subject to section 351 of the Public Health Service Act.

- "Acceptance and Filing Reviews for Premarket Approval Applications (PMAs)" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM313368.pdf)

Further information on device regulation can be found at "Device Advice" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/default.htm).

FDA recognizes that biodosimetry device sponsors face unique challenges in attempting to meet premarket clearance or approval requirements for their devices. For example, ethical considerations limit the availability of appropriate clinical samples for biodosimetry device performance studies, leading to the expected use of animal models as surrogates for clinical validation purposes. As such, study designs that are not typically recommended for other IVD submissions may be appropriate for radiation biodosimetry devices. The intent of this guidance document is to provide clarity to both industry and FDA staff on the performance data that should be included in a premarket submission in order to help sponsors design studies, and to facilitate our review process. Due to the device's novelty, the pre-submission process is an encouraged component of any radiation biodosimetry device development plan. Information on the pre-submission process can be found in the guidance document entitled "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM311176.pdf).

III. Scope

This draft guidance only applies to validation of biodosimetry devices intended to be used to assess exposure in non-therapeutic or accidental scenarios (e.g., a deliberate attack, such as use of an improvised nuclear device, or a natural disaster). This guidance neither applies to devices that assess deliberate radiation dosing that may occur in the course of medical treatment nor to devices that measure effects from long term exposure. In addition, dosimeters, which are devices that detect radiation exposure on a physical substrate rather than through a biological response and are worn by people who might be exposed to radiation during the course of their normal work (such as film badges), are not addressed in this guidance document. Finally, biological assays that might be used to detect the presence of ingested radioisotopes in sputum or urine are not considered in this guidance document.

IV. Policy

A. Benefit-Risk Analysis

A source of significant risk to patient health associated with radiation biodosimetry devices is when a failure of the device to perform as indicated leads to either deficient or inaccurate results or the incorrect interpretation of these results. Inappropriate or incorrect use of radiation biodosimetry devices may be an additional risk. These potential risks may then

lead to incorrect patient management decisions. For instance, a false positive result or overestimation of exposure in an emergency scenario with adequate resources could lead to unnecessary or inappropriate treatment for ARS. Alternatively, an overestimation of exposure in a resource poor mass exposure scenario may result in the patient being inappropriately placed in an expectant category and given only palliative care when treatment could be life-saving. Likewise, from a public health standpoint, a false positive result in a resource poor response area could lead to a misallocation of resources. However, a false negative result or underestimation of exposure could lead to failure to provide treatment or incorrect patient management which may be lethal. Thus it is essential to appropriately balance the benefits and risks of false negative and false positive results.

Current laboratory methods for determining absorbed radiation doses utilize cell based assays that are accurate, but take several days to complete. Therefore, part of evaluating the benefit/risk considerations for new radiation biodosimetry systems will be evaluating performance differences along with the overall time to result or throughput of the system (e.g., decreased accuracy may be acceptable for a device that provides a more rapid result in the context of triage). The guidance document entitled "Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approvals and De Novo Classifications" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm267829.htm) provides information on FDA benefit-risk determinations. Premarket submissions should include a discussion of the potential benefits and risks associated with the device in light of the biological radiation response pathway that is being assessed, the analytical strengths and weaknesses of the technology, and the clinical information that is available demonstrating device effectiveness.

B. Device Description and Specifying the Intended Use

All components of your radiation biodosimeter system necessary to achieve the claimed functionality in the intended use statement should be listed in the device description.

The intended use statement should specify 1) the nature of the analyte (e.g., RNA, DNA, or protein), 2) specimen types in which testing may be performed (e.g., blood, urine, or saliva), and 3) the specific population(s) for which the test is intended (e.g., pediatrics, general population).

The intended use statement should also explain whether the test is qualitative or quantitative and include any specific conditions of use. The intended use statement of most radiation biodosimetry devices should also explicitly advise that results need to be considered in combination with other appropriate clinical signs and symptoms as well as radiation dispersal monitoring.

Additional specific elements that should be considered for radiation biodosimetry devices include the following:

1. The stage of response for which your device is intended

Biodosimetry devices may be designed for preliminary triage during a mass exposure event. For these types of devices, the intended use should specify the throughput capabilities, and time to result of the device and the decision making cut-points assessed. Alternatively, biodosimetry devices may be designed for dose level confirmation and medical management at later stages of a mass exposure scenario or in situations where only a small number of people need to be assessed. For these types of devices, the assay analytical range and specific clinical indicators of health status should be part of the intended use statement.

2. Appropriate time-frames for testing

Because many common biomarkers of radiation exposure display defined kinetics, during which they become detectable, remain stable, and then disappear from the matrix being examined, it is important to specify the time-frame in which the device is designed to function, beginning from time of exposure. This should include both the beginning and end of the acceptable testing window (e.g., from 30 minutes to 48 hours post-exposure).

3. Assay limitations

Validation of radiation biodosimetry devices may be incomplete due to a lack of samples from the intended use population. Therefore, limitation statements may be needed to minimize risk of over-reliance on biodosimeter results when the real world situation in which the device is being used does not mirror the scenarios tested. For instance, if validation testing was only performed on certain populations (e.g., not tested in pediatrics), for specific radiation types (e.g., exposure to gamma source only), or on specimens or subjects exposed to limited doses and dose rates that may not reflect the final situation of use, these limitations should be captured in the labeling. Other situations that may cause inaccurate biodosimeter results (e.g., combined injury such as radiation plus physical trauma) should be taken into account when drafting an appropriate intended use statement. However, labeling limitations cannot generally be relied upon to justify missing validation studies using the intended use population. You should make every attempt to capture the intended use population. If you determine that you are unable to perform validation using the intended use population, you should provide a justification along with a detailed description of the due diligence activities you performed to support your validation approach.

C. Establishing Performance Characteristics – Analytical Validation Studies

The most significant benefit that radiation biodosimetry has over standard dosimetry is that biodosimetry takes into account the natural patient biological variability in radiation response, though dosimetry tools may provide a more accurate representation of the actual radiation dose delivered. For example, while a standard dosimeter may give an accurate representation of dose, standard dosimetry devices will be unable to differentiate a patient

who is radiation sensitive (presumably needing a higher level of medical intervention) from a patient who is radiation resistant. As such, the development of radiation biodosimetry techniques will be a powerful tool in personalizing therapeutic responses to radiation exposure, representing a significant benefit to public health. However, because biological response to a radiation dose is measured by biodosimetry, the measurement is confounded by patient-to-patient variation in radiation resistance making simple dose/response correlations difficult. As such, a well-documented explanation of the relevant biological pathways should be provided to FDA in order to justify a lack of correlation to an accuracy standard attributable to natural biological response. Such information can be obtained through peer-reviewed literature as well as bench testing. Well-controlled analytical studies should be provided to FDA to establish device performance across the entire analytical range of the device in a defined sample subset. This analytical performance information will be critical to substantiate intended use claims of a radiation biodosimetry device.

As discussed in section II, radiation biodosimetry devices generally will be IVDs. Accordingly premarket submissions will generally be reviewed by the Office of *In Vitro* Diagnostics and Radiological Health (OIR). This section provides specific recommendations to facilitate planning your validation studies. You are encouraged to consider relevant guidance documents for general information on what to include in submissions and the types of studies that may be expected. The following two documents may be useful in preparing premarket submissions:

- "In Vitro Diagnostic (IVD) Device Studies – Frequently Asked Questions" (http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm078309.htm)
- "eCopy Program for Medical Device Submissions" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM313794.pdf)

Additionally, it is beyond the scope of this document to consider the broad range of analytical characteristics specific to each of the many and varied technologies that may be employed to develop radiation biodosimeters. Therefore, you are encouraged to examine guidance documents that might be applicable to the type of technology your device employs to identify the types of analytical characteristics that might be appropriate to demonstrate for your device. For example, if a biodosimetry instrument utilizes a genetic test for heritable markers, you might consider consulting the "Guidance on Pharmacogenetic Tests and Genetic Tests for Heritable Markers" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071075.pdf) to identify the types of analytical characteristics that might be appropriate to demonstrate for your device.

Finally, appropriate standards documents drafted by the Clinical and Laboratory Standards Institute (CLSI) are additional helpful resources that might provide details on specific analytical performance testing appropriate for your device. Specific standards documents which might be applicable will be referenced throughout this section. A list of FDA

recognized standards can be found at the following website: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfStandards/search.cfm.

1. Sample availability

We recognize the difficulty in acquiring appropriate clinical samples for use during device validation. Two possible sources of clinical samples include site-limited radiation therapy patients and total body irradiation (TBI) patients. However, these specimens may be difficult to obtain in sufficient amounts. Another challenge with these sources is that the therapeutic protocol restricts the dose exposures available (only low dose fractionated exposures may be readily obtained). In the absence of sufficient or appropriate clinical samples, contrived samples may be used to supplement clinical samples for analytical performance testing, but only as a supplement to, and not as a substitute for, clinical samples.

Samples may be contrived by *ex vivo* irradiation of the appropriate matrix, spiking the analyte of interest into the appropriate matrix, or through the use of animal-derived specimens, as appropriate. In some instances, control material may be used for the purposes of analytical performance testing. Premarket submissions should include a scientific justification for contrived sample utilization, a description of how contrived samples were generated and validated for testing, and a description of how the results obtained translate to the clinical setting. You are encouraged to discuss the most appropriate sample type for testing with FDA prior to designing your analytical validation studies. The proportion of samples that may be contrived for analytical validation testing will depend on both the technological characteristics of the device and the abundance of appropriate samples. You should be thorough in attempting to obtain and use appropriate clinical samples to demonstrate performance.

2. Specimen collection and handling

The quality and quantity of an extracted analyte can be affected by multiple factors such as specimen source, collection method, and handling (e.g., transport, storage time, and temperature). Therefore, premarket submissions should include performance validation data to establish that the specimen collection and transport system employed by the device provides an adequate and appropriate yield of the analyte being detected by the assay (e.g., DNA or RNA from blood or tissue). Testing should also demonstrate that the device maintains acceptable performance under all the various specimen handling conditions claimed in the product labeling.

Specimen stability should be addressed in the radiation biodosimetry premarket submission. However, we recognize there may be different analytical and clinical performance needs depending on whether the device is intended for initial triage in a field setting, for radiation exposure confirmation, or for dose refinement in a clinical laboratory setting.

For example, if the device is intended to be used in a field triage environment, CDRH would evaluate whether the device is both sufficiently robust to withstand environmental impacts

and appropriately user friendly for use by the intended user. Therefore, for biodosimetry devices intended for field triage use, performance testing should include performance testing to demonstrate the robustness of the device, and where relevant, performance testing to demonstrate the device's specimen collection and transport performance characteristics. By contrast, performance evaluation of devices intended for dose refinement may focus more on performance testing to demonstrate the device's measurement precision performance characteristics. In either case, specimen shipping stability performance testing will be critical if the testing is intended to take place far from where the patient samples are obtained. The acceptance criteria for all specimen stability parameters should be clearly indicated and justified in terms of the intended use environment as indicated in the labeling.

3. Accuracy

In order to demonstrate analytical accuracy, the device's measurement of the biological response to radiation is compared to the physical calculated dose delivered. Therefore, the accuracy of the delivered dose is crucial, and the protocol for designing proper telemetry should be included in all study protocols. As discussed above, we expect that because the biodosimetry output will confounded by the biological response to radiation, there will be inter-individual variations which may complicate the correlation of the output to the accuracy standard.

In the case of human clinical samples, since the exposure rates and overall dose will be dictated by the therapeutic protocol, there will be no need to justify the doses and dose rates used. Nevertheless, submissions should include information on the radiation source, dose delivered, dose rate, and the time intervals at which device testing was performed. When animal studies are being used to supplement human clinical samples, a scientific justification for the dose rates and doses delivered should be included in the submission, in addition to information on the radiation source, dose delivered, dose rate and time intervals. Animal studies should be designed to "bridge" between animal and human biological response with the human radiation protocol duplicated in the animal study, while other animal studies should supply information on test performance using doses/dose rates that cannot be ethically obtained in human clinical studies.

Statistical analysis plans including acceptance criteria should be developed around accuracy studies, and be appropriate for the device output (qualitative or quantitative). Please refer to Appendix A in section V of this document for more information on statistical considerations for study designs. All accuracy studies should be designed to demonstrate device performance at the relevant clinical decision making points relevant to the intended use of your device (i.e., triage or medical management) and contain pre-specified success criteria. The rationale for your acceptance criteria should be provided and clinically justified.

The reference method used to evaluate accuracy of the biodosimetry device will depend on your device's intended use. For example, if the device is intended for use in both triage and confirmation, a reference method currently used for small-scale radiation exposures, such as the dicentric chromosome assay, may be appropriate. However, chromosome-based assays

may not be an appropriate comparator for devices designed for triage due to the time needed to perform this reference method.

4. Analytical range

If a radiation biodosimetry device reports numerical (quantitative) results, submissions should include studies establishing the analytical range of your device. Analytical range studies should be designed to substantiate Limit of Detection (LoD), Limit of Blank (LoB), Limit of Quantitation (LoQ) and the linear range claims of the assay. The clinically relevant range of exposure is 0-10 Gy and we recommend developing test protocols to demonstrate the limitations of your assay outside the reportable range (both above and below). For instance, if high exposure levels can cause the assay to incorrectly report lower values (i.e., the hook effect), then the limitations section of the labeling would include this information. Please refer to the CLSI document EP6-A, "Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach," and EP17-A2, "Protocols for Determination of Limits of Detection and Limits of Quantitation" to assist in the design of such studies. Samples around critical decision making cut-points should also be included in analytical range validation studies.

5. Interference

Interfering substances may confound assay outputs. As such, radiation biodosimetry device submissions should include the results of appropriate interference studies (see CLSI Document EP7-A2 "Interference Testing in Clinical Chemistry"). Interference studies may be specific to the matrix being examined. For example, if the assay uses blood, interference from hemoglobin, bilirubin, and lipids should be examined to mimic grossly hemolytic, icteric and lipemic samples. Likewise common drugs known or expected to interfere, or drugs that are expected to be administered in a mass radiation exposure scenario, should be tested for assay interference. See the document entitled "Planning Guidance for Response to a Nuclear Detonation" (http://www.remm.nlm.gov/PlanningGuidanceNuclearDetonation.pdf) for more information on the Federal government response plans for radiation disaster scenarios and the applicable drugs and treatments that will be recommended in such a scenario. (Note this website is not controlled by FDA. The content of the website was last verified on December 18, 2014.) In addition, you should consider how biological responses may interfere with your assay and develop corresponding risk mitigations necessary to provide a reasonable assurance of safety and effectiveness (e.g., to limit the interfering effects of the underlying biological response). This information may be appropriate to include in the warnings and limitations section in the labeling. Ultimately the decision of what compounds to test for assay interference should be based on applying scientific reasoning to the technological characteristics of your device and the major biological pathways being interrogated. For instance, if a radiation biodosimeter incorporates expression profiles from a pro-inflammatory pathway, interference from anti-inflammatory drugs should be examined and assay effectiveness should be tested in normal volunteers with possible confounding diseases such as arthritis and other inflammatory conditions.

Alternatively, animal models may be used to evaluate these potential confounders (see section IV(D) below).

6. Other analytical testing protocols

Analytical testing protocols, in addition to those mentioned specifically above, should be included in radiation biodosimetry submissions as applicable given the intended use, output, and technology. Reproducibility generally should be demonstrated at a minimum of three sites (of which at least one should be in the United States) and submissions should include an analysis of site-to-site, operator-to-operator, instrument-to-instrument, and kit lot-to-kit lot reproducibility as applicable. The limits of detection, quantitation, and blank, as relevant, may be demonstrated in a separate study or in combination with the analytical range study. The stability of kit reagents should be performed in real-time for product expiry dating, and should also be examined in shipping simulation studies. As noted above, studies to support technology-specific validation and special controls may apply, such as for molecular assays.

In addition to CLSI documents indicated elsewhere in this document, the following may be useful in understanding how such analytical performance studies are typically designed:

- EP5-A2, "Evaluation of Precision Performance of Quantitative Measurement Methods"
- EP9-A2, "Method Comparison and Bias Estimation Using Patient Samples"
- EP12-P, "User Protocol for Evaluation of Qualitative Test Performance"
- EP17-A2, "Protocols for Determination of Limits of Detection and Limits of Quantitation"
- EP25-A, "Evaluation of Stability of *In Vitro* Diagnostic Reagents"

7. Controls and calibrators

The design of radiation biodosimetry devices should incorporate the use of on-board or external controls as appropriate. Any recommended Quality Control (QC) procedures and acceptance criteria used during the analytical and clinical validation of your device should be included in the instructions for use. A description of the control material and its recommended use should be submitted along with the assay for premarket clearance or approval. Information that should be provided in the premarket submission includes control performance, value assignment, and reagent stability as outlined in guidance entitled "Assayed and Unassayed Quality Control Material" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm079179.htm). Similarly, if external calibrators are required for the assay system, they should also be submitted with the assay. You are encouraged to consult the guidance entitled "Abbreviated 510(k) Submissions for In Vitro Diagnostic Calibrators" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm092801.pdf).

8. Instrumentation and software

If your radiation biodosimetry device utilizes software, you should submit the information listed in the guidances entitled "Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm089543.htm) ("Software Premarket Submissions Guidance" for the duration of this document) and "Guidance for Off-the-Shelf Software Use in Medical Devices" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm073778.htm). The information you should submit is determined by the "level of concern," which is related to the risks associated with software failure, as explained in the Software Premarket Submissions Guidance.

D. Establishing Performance Characteristics – Animal Studies as Surrogates for Clinical Validation

Traditionally, animal data are not used in IVD submissions. IVD sponsors are expected to obtain specimens from the intended use population either through prospective collection from a trial or to use properly archived surplus (excess) specimens to demonstrate test performance when prospective studies are not feasible. However, in the case of radiation biodosimetry devices, we acknowledge that appropriate human specimens may not be available. Therefore, under the following conditions it may be appropriate to use animal model data to supplement human clinical samples to demonstrate device performance:

- The analyte(s) being detected is not stable in archived specimens;
- A diligent search of available specimen banks has failed to yield adequate samples for testing; or
- A prospective trial is either unethical, or prospective trials that may be ethically performed will not yield a sample set adequate to demonstrate assay performance over the analytical range of the device.

1. Defining an appropriate animal model

If animal model data is included in a biodosimetry device premarket submission, the submission should also include an appropriate justification for the animal model or models chosen. You should provide evidence that the model is an appropriate substitute for human specimens. In particular, establishing that there is high homology in the analyte(s) being assessed and that the animal model displays similar responses to radiation exposure in the biological pathways being interrogated is important. While rodent models may be appropriate for early proof of concept studies, they are not typically adequate for demonstrations of device effectiveness. Primate or porcine models, in which radiation biological response pathways are well understood, may be the best option for effectiveness studies. You should ensure that an equal distribution of genders is included in pre-clinical testing and consider including animals at various age ranges (e.g., juvenile, adult, elderly) so

age and gender-related differences in radiation response can be assessed for the analytes included in the biodosimetry testing.

As stated above, for biodosimetry device effectiveness studies, animal models should supplement and not be a substitute for human clinical samples. Therefore, multiple animal models generally are not required to demonstrate the biodosimeter's effectiveness. However, multiple animal models may be necessary when a single appropriate animal model cannot be identified for all analytes being assessed by the device.

Please note that all device effectiveness studies using animal models must comply with good laboratory practice for nonclinical laboratory studies regulations as described in 21 CFR Part 58, including 21 CFR 58.90. Further, FDA recommends that you follow the Animal Welfare Act, PHS policy, and their applicable regulations and statutes. FDA believes that following these regulations enhances the opportunity and intensity of observations and can potentially result in other useful findings for the investigators (see Ref. 1-4).

2. Effect of "supportive care" on device output

In some cases, animal housing and supportive care conditions might influence radiation biodosimetry assay results. For instance, analyte expression might be influenced by differences in diet or lifestyle patterns of laboratory animals as opposed to that of the general human population. You should make an attempt to address the effect of confounding factors that may be associated with aspects of animal care that will not be reflective of the intended use population. Additionally, to avoid variability caused by changes in radiation response due to circadian rhythm patterns, you should deliver radiation at the same time of day (e.g., a.m. vs. p.m.) to all animal models in pre-clinical testing.

You also may want to consider providing animal models with the same supportive care that is expected to be provided to people in a radiation mass exposure scenario, such as antibiotics and fluids, to determine if these types of medical interventions alter the biological responses being assessed by biodosimetry. See the document entitled "Planning Guidance for Response to a Nuclear Detonation" (http://www.remm.nlm.gov/PlanningGuidanceNuclearDetonation.pdf) for more information (Note this website is not controlled by FDA. The content of the website was last verified on December 18, 2014). For example, if neutrophil numbers are part of a biodosimetry algorithm, it should be understood how G-CSF treatments affect resulting exposure level estimations. This information may be critical to ensuring appropriate labeling is provided for the assay.

3. Bridging animal data to human data

When animal model data are necessary to demonstrate the effectiveness of a radiation biodosimeter across the reportable analytical assay range, then two types of studies should be provided. First, a set of experiments should be designed to bridge between the animal results and the available human clinical information. For instance, if available human samples were

derived from whole body exposure delivered in 2 Gy fractions over the course of a week, then an animal study should be performed to mirror this exposure pattern to demonstrate how animal results and human data are reflective of each other. Once an appropriate demonstration has shown that animal results can be bridged to human clinical experience, then studies in animals may be performed to address device performance at radiation doses, dose-rates, time-courses, and sources that cannot be examined in human clinical studies.

Animal studies may also be used to address situations that may not be easily addressed using human clinical studies. For instance, you should consider testing with commonly used drugs such as anticholesterol, antihypertension, diabetic drugs, and other common drugs if there is evidence to suggest that they may interfere with the analyte being evaluated (e.g., a particular metabolic pathway CYP450, growth factors). Other situations that may confound assay results such as combined injury can also be examined in pre-clinical models.

Animal studies should be designed to use the fewest possible animals to demonstrate statistical significance. Because these studies will be critical to demonstrate the performance of a radiation biodosimeter and because alternative study designs will be needed to minimize animal numbers, it is encouraged that study protocols be submitted to us prior to the onset of testing in order to gain Agency concurrence on study parameters and statistical analysis plans. Please refer to the appendix in section V of this document for more information on statistical considerations for radiation biodosimetry study designs.

E. Establishing Performance Characteristics – Clinical Validation Studies with Human Samples

FDA expects that radiation biodosimeter safety and effectiveness will be established in a clinical study that contains appropriate human samples. As discussed above, non-clinical models may be used to supplement human clinical data by providing data on doses, dose rates, and radiation sources that cannot be ethically obtained in a human clinical study. However, manufacturers should strive to perform validation studies using the intended use population of the device. These pivotal studies should be designed with an appropriate statistical analysis plan in place with pre-defined acceptance criteria. You are encouraged to refer to Appendix A in section V of this document and the guidance entitled "Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071148.htm) for more information.

1. Prospective clinical studies

Samples may be collected prospectively for biodosimeter clinical testing in the context of radiation exposure for therapeutic purposes. For instance, patients may consent to provide blood, tissue, or other relevant samples during the course of standard radiation therapy or in the context of radiation therapy clinical trials. Information should be captured on the dose, dose rate, and source of exposure for all patients included in prospective testing. In addition,

basic demographic information should be collected, including patient age, gender, and race. Whenever possible, information on the patient's clinical condition and medications should also be collected.

Prospective IVD clinical studies that do not require an invasive sampling procedure and for which the test results are not used to support patient management are generally considered to meet the requirements under 21 CFR 812.2(c)(3) to be exempt from the Investigational Device Exemption requirements in part 812 with the exception of 21 CFR section 812.119. If, however, you have concerns about the risk classification of your prospective clinical study, you should submit a risk-determination pre-submission as outlined in the "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program and Meetings with Food and Drug Administration Staff" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM311176.pdf).

Because patient numbers may be limited, you should consider a study design that collects multiple samples from patients over the course of their therapy. For example, consider designing studies to assess biodosimeter performance over the range of times post-exposure that you want to capture in labeling (such as 24 hours–7 days post-exposure). Samples should also be collected prior to radiation exposure whenever possible as control specimens.

We acknowledge that there are a number of challenges associated with prospectively collecting adequate clinical samples for device effectiveness testing. Therefore, it is recommended that you use the pre-submission process to discuss clinical study design and implementation prior to the onset of testing. Pre-submissions should ideally include the clinical validation study protocol, statistical analysis plan, and a description of sample acquisition strategies.

2. Retrospective clinical studies

If the analyte or analytes being assessed by a biodosimetry device are suitably stable in the relevant test matrix, then you are encouraged to utilize appropriately banked retrospective samples to demonstrate the clinical performance of your biodosimetry device. For more information, please refer to the guidance entitled "Informed Consent for *In Vitro* Diagnostic Device Studies Using Leftover Human Specimens that are Not Individually Identifiable" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm078384.htm). Appropriate retrospective samples should have documentation on basic patient demographics (but not personally identifiable information) and basic information on the radiation exposure profile.

3. Normal control samples

Normal (unirradiated) control samples may be used both to assess the normal range of analyte expression and to create contrived samples by spiking the analyte of interest into an appropriate matrix. Assessment of the normal range of analyte expression should include a

suitable number of normal control samples to determine if biodosimeter analyte expression is affected by subject age, race, gender, and common health conditions (e.g., obesity, diabetes, or arthritis). Normal control samples can be obtained prospectively or from appropriate sample banks. All normal control samples should be collected as intended for your device. For instance, if blood will be collected in EDTA tubes for use in your biodosimeter, then normal control samples should be collected in EDTA tubes. As above, basic demographic information and information on health conditions should be collected with normal control samples.

4. Limitations of clinical studies

We acknowledge that there may be significant limitations in the interpretation of clinical studies for radiation biodosimetry submissions. For instance, therapeutic dose rates will not be reflective of the expected dose rates that would be experienced by someone in a radiological disaster. Additionally, clinical studies may not be reflective of all possible types of radiological disasters for which the biodosimeter is intended. Thus, in addition to the animal studies and analytical studies discussed above, biodosimetry submissions should include a discussion of the limitations of the clinical study, and how analytical studies, animal studies, or a combination of analytical and animal studies have been used to mitigate these limitations in order to demonstrate a reasonable assurance of the safety and effectiveness of the device for its intended use.

F. Labeling

The labeling of a radiation biodosimetry device includes the instructions for use, package inserts, and any outer box or container labels for the device itself, reagents, and control materials, as applicable. The following references will be useful in developing clear and complete labeling for your device.

- The guidance entitled "Guidance on Medical Device Patient Labeling" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm070782.htm)
- CLSI document GP-14 "Labeling of Home-Use In Vitro Testing Products"
- Labeling Requirements – In Vitro Diagnostic Devices (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/DeviceLabeling/InVitroDiagnosticDeviceLabelingRequirements/default.htm)

In addition to the references above and the information already provided in this guidance document, the following should be considered when developing labeling for radiation biodosimetry devices that complies with the labeling requirements outlined in 21 CFR Parts 801 and 809:

1. Instructions for use

For radiation biodosimeters intended for use by lay persons, information required by 21 CFR 809.10(b) should be described in a manner that lay users can understand. Detailed technical

information (e.g., scientific principles of test procedure or statistical analysis of data) may be presented in a separate section followed by clarifying statements appropriate for lay users. The following should also be taken into account when drafting appropriate instructions for use for radiation biodosimetry devices.

- The labeling must provide instructions for specimen collection and preparation (see 21 CFR 809.10(b)(7)). Instructions should be drafted with the intended end user in mind. For example, consider whether a trained healthcare provider will be collecting the sample or if the patient will be instructed to do so.
- The labeling must provide a step-by-step outline of recommended procedures and operating instructions for the instrument (see 21 CFR 809.10(b)(8) and 21 CFR 809.10(b)(6)(v)). Ideally, numbering rather than bullet points should be used for clarity.
- Labeling must describe details of calibration and of quality control procedures (see 21 CFR 809.10(b)(8)(v) and 21 CFR 809.10(b)(8)(vi)). These instructions are to help ensure optimal performance of the system. This section should include recommendations for how and when to perform quality control checks and instructions for what to do if the control material values are not within the allowable ranges.

2. Limitations

Labeling must include a statement of limitations of the procedure, including known extrinsic factors or interfering substances affecting results (see 21 CFR 809.10(b)(10)). You should also include testing conditions that may cause clinically significant errors due to bias or imprecision (e.g., combined injury, high dose rates, or alternative sources of radiation). You should also note all known contraindications. This section should thoroughly describe the situations of use that were not examined in performance testing of the device. For instance, if device performance was not assessed for performance using neutron radiation sources, then this information should be included in the labeling.

3. Interpretation of results

Labeling must include expected values for your device (see 21 CFR 809.10(b)(11)). We recommend that the expected values be portrayed in terms of expected values for non-irradiated patients, and around the clinical decision making cut-points that were evaluated for your device (e.g., 2 Gy and 10 Gy). If the results are qualitative, you should explain how to interpret positive and negative results, including their clinical significance. If the results are quantitative, you should explain how numerical results correlate with expected values and the clinical significance of outputs. A statement should be included to interpret results in the context of other clinical signs and symptoms as well as any known dosimetric or radiation dispersal data associated with the patient's location.

You should also provide directions for the interpretation of the results of controls (performance monitors) and provide a statement that if controls do not perform as expected,

assay results are invalid. Instructions should also be provided for any situation in which the end user should repeat a test.

4. Performance characteristics

Labeling must include specific performance characteristics of the device (see 21 CFR 809.10(b)(12)). All studies, including bench testing, animal testing, and clinical studies should be summarized in the package insert of the assay. Performance data should be presented clearly and accurately, ideally in both graphical and text formats. Clinical information that was obtained through animal studies alone (with no human clinical supporting data) should be specifically highlighted with a disclaimer that the performance of the assay has not been evaluated in clinical samples under these specific conditions.

G. CLIA Categorization

As discussed above, some radiation biodosimeters may be designed for use in a clinical laboratory, while others may be designed to be used in a professional healthcare facility such as a hospital. Radiation biodosimeters intended to be used in initial triage may be designed to be performed outside of professional healthcare facilities or clinical laboratories by laypersons. The location where a device is intended to be used (e.g., clinical laboratory, professional healthcare facility, or home use) impacts the CLIA categorization of the device, and informs the kind of information that will need to be included in submissions to allow CLIA categorization to be completed. A document that may be of particular interest is the guidance entitled "Design Considerations for Devices Intended for Home Use" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM331681.pdf). This guidance provides suggestions to assist manufacturers in designing and developing home use devices that comply with applicable standards of safety and effectiveness and other regulatory requirements.

CLIA (codified at 42 U.S.C. 263a) established uniform quality standards for all laboratory testing to ensure the accuracy, reliability and timeliness of patient test results throughout the United States. Under CLIA, laboratory tests are categorized according to complexity in order to determine what level of certification, if any, a laboratory or other user will be required to have in order to perform human diagnostic testing. As the complexity of a test increases, the number of entities certified to use it becomes more limited. Waived tests, which are simple tests, may be used by a variety of users, including inexperienced users. More complex tests known as moderate complexity tests may be performed in laboratories certified as moderate or high complexity (these can include health care provider offices). High complexity tests are those that are either difficult to perform or difficult to interpret, and may be performed by only specific clinical laboratories in the United States certified to perform high complexity testing.

Since 2000, CDRH has been responsible for categorizing commercially marketed IVDs under CLIA. CDRH determines the CLIA categorization of IVDs at the time of premarket submission. Thus, information pertaining to CLIA categorization should be included with

the initial premarket submission. In particular, if your device is intended to be used in a field triage environment, you should design studies to demonstrate the performance of your device, including specific human factors, to verify performance by lay personnel in a non-laboratory setting. You are encouraged to review the guidance entitled "Administrative Procedures for CLIA Categorization" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm070762.htm) for information on how IVDs are categorized, and the information needed for CLIA waiver applications. In addition, the guidance entitled, "Design Considerations for Devices Intended for Home Use" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM331681.pdf) provides suggestions to assist manufacturers in designing and developing home use devices that comply with applicable standards of safety and effectiveness and other regulatory requirements.

V. Appendix A: Statistical Considerations for Radiation Biodosimetry Devices

This appendix includes some statistical considerations for radiation biodosimetry devices. Further statistical considerations for diagnostic devices that may be applicable to radiation biodosimetry devices are comprehensively discussed in the guidance entitled "Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm071148.htm)and further concepts and principles related to designing medical device studies are discussed in the guidance entitled "Design Considerations for Pivotal Clinical Investigations for Medical Devices" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM373766.pdf). In addition, as discussed above, it is recommended that you use the pre-submission process to discuss statistical analysis plans and methods before initiating any data collection for your clinical, pre-clinical and analytical studies.

A. Independent Validation

Validation of the performance of a radiation biodosimetry device should be conducted on subjects, which includes specimens from subjects that are independent of those used during device development. Before a validation study is conducted, all of the specifications for the assay (e.g., algorithm, probes, manufacturing methods, cut-off values) should be "locked down" (in the final version form). Generally, changes subsequent to validation would create the need for additional analytical and clinical validation studies separate from those already performed.

B. Study Design

When evaluating the design of a study intended to establish the safety and effectiveness of a radiation biodosimetry device, a main consideration is whether the design could introduce non-negligible bias into the estimation of device performance. In a biased study design, estimates of device performance will tend to deviate systematically from the true performance of the device in the intended use population, regardless of the size of the study. For example, bias may be introduced in the selection of subjects, which includes specimens from subjects, study conduct, and mechanisms of data analysis, and may also arise from missing data. Understanding potential sources of bias and how to avoid or minimize them during the design of your study is essential. Some strategies for this purpose are described in the guidance entitled "Design Considerations for Pivotal Clinical Investigations for Medical Devices" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM373766.pdf). You may also refer to the extensive literature with comprehensive discussions on sources of bias in diagnostic devices studies.[2]

C. Study Integrity (Blinding)

To avoid bias in the device result, the user of the radiation biodosimetry device should be unaware of (i.e., blinded to) the actual level of radiation exposure (e.g., the administered dose or the level as measured by the clinical reference standard). Likewise, to avoid bias in the reference measurement, the user of the reference standard should be unaware of the device result. In general, the user should be unaware of any results from other diagnostic evaluations, and vice versa.

D. Precision of Estimation

The sampling variability or the precision of estimation is controlled by the sample size of the study and is another key consideration when evaluating a study design and study results. With a larger sample size, an estimate of performance is subject to less sampling variability. Uncertainty of the estimation is thus reduced. The estimate becomes less imprecise, leading to a narrower confidence interval of likely values for the true performance. If you have more than one study, the precision of estimation might be increased by a careful pre-planned analysis of combined studies, if appropriate.

E. Study Analysis

The protocol for a radiation biodosimeter study should include a Statistical Analysis Plan (SAP). The SAP is used to interpret the study data in support of the safety and effectiveness of the device for its intended use. The SAP should be pre-specified and provided in enough detail to permit FDA review. The analysis plan should define the performance measures

[2] E.g., Pepe M, The Statistical Evaluation of Medical Tests for Classification and Prediction, Oxford, 2003; Zhou XH, Obuchowski NA, and McClish DK *Statistical Methods in Diagnostic Medicine*, 2nd ed, Wiley 2011; Begg C. Biases in the Assessment of Diagnostic Tests, *Stat Med* 1987; 6: 411-423.

(e.g., specificity, sensitivity) to be evaluated. Any success criteria on the performance measures should be clearly pre-defined using descriptive and mathematical statements (e.g., the null and alternative hypotheses for a significance test, the estimator for a performance measure, the method for deriving a confidence interval, etc.). Unplanned post-hoc analyses are discouraged as primary evidence of safety and effectiveness. Post-hoc analyses can inflate the study-wise type I error rate (probability of false statistical significance) and therefore are generally considered exploratory rather than confirmatory evidence of safety and effectiveness.

In particular, the SAP should describe how sample size for the study was determined. Sample size determination should be consistent with the pre-planned statistical analyses of the study, especially the primary analyses. Assumptions underlying the statistical power of the study to demonstrate a performance claim should be provided in detail.

The SAP should include a plan for dealing with device results that are considered uninterpretable, invalid, indeterminate, equivocal, or missing, and samples or specimens that are unavailable or unevaluable. For example, the plan may include reporting the number and proportion of subjects without a valid device result by the reason a valid result was not obtained. If repeated application (after the first reading) of the device on a subject or specimen is not intended, is not possible, or would not be helpful (e.g., the device result will always be uninterpretable), then uninterpretable device results, for example, could be treated as a separate category for the purpose of analysis. If repeat measurement of subjects or specimens is possible and appropriate, then imputation of missing device results can sometimes aid the statistical analysis and interpretation of study data. Further comments with respect to missing data are in section V(E)(4) below.

If the design of your study is adaptive, its adaptive features should be pre-planned (i.e., the study should be adaptive *by design*). Before conducting an adaptively designed study, its operating characteristics (e.g., type 1 error rate, power) and its potential for introducing operational bias into the study (due to the adaptive features) should be evaluated. Some examples of adaptations based on interim analysis include stopping the study early for futility or success, re-estimating sample size, and changing a hypothesis. Monitoring a study until it has a pre-defined number of subjects with or without a condition is an adaptive design feature which does not ordinarily require special consideration. The literature on adaptive design for diagnostic studies is unfortunately scant, but some references are available.[3]

1. Quantitative, continuous, or semi-quantitative output

Statistical analyses and methods depend on the type of results provided by the radiation biodosimetry device. For devices providing a quantitative measurement, the bias and

[3] Gerke O, Hoilund-Carlsen PF, Poulsen MH, Vach W. Interim analyses in diagnostic versus treatment studies: Differences and similarities. *Am J Nucl Med Mol Imaging* 2012; 2(3): 344–52; Tang LL, and Liu A. Sample size recalculation in sequential diagnostic trials, *Biostatistics* 2010; 11(1): 151-163.

imprecision of the measurement should be evaluated along with other performance characteristics. For example, Bland-Altman methodology can be used to compare the level of radiation exposure predicted by the device with the actual level of exposure (actual dose or reference level) over the measuring interval.[4] Comparison of device and reference results near clinical decision making cut-points is essential. The bias of the device result should be estimated near these decision points. Please refer to the CLSI document EP09-A3, "Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline-Third Edition," and EP5-A2, "Evaluating of Precision Performance of Quantitative Methods; Approved Guideline-Second Edition."

Receiver Operating Characteristic (ROC) analysis might also be useful to evaluate the diagnostic accuracy of a radiation biodosimetry device reporting a quantitative, continuous, or semi-quantitative result, or reporting a qualitative result derived from an underlying value that is quantitative, continuous, or semi-quantitative. ROC analysis evaluates the overall ability of the device to discriminate between subjects with and without a condition of interest. On an ROC plot, the false positive and true positive fractions (1 – specificity, sensitivity) are plotted for each possible cut-off in the value as it is varied across the entire range of observed values, resulting in an ROC "curve." An advantage of ROC analysis can be that the plot will display the estimated sensitivity and specificity of the device throughout a range of clinical decision making points. The area under the ROC curve (AUC) is a global measure of device discrimination performance, with AUC values of 0.5 and 1.0 indicating random and perfect discrimination, respectively. Please refer to the CLSI document EP24-A2, "Assessment of the Diagnostic Accuracy of Laboratory Tests Using Receiver Operating Characteristic Curves; Approved Guideline-Second Edition," as well as the following references.[5] If the condition of interest is not binary (e.g., dose or level of radiation exposure), generalizations of AUC exist for evaluating the discrimination ability of medical tests.[6]

2. Qualitative output

For devices that are designed to provide a qualitative output around a clinical decision making cut-point, device evaluation should include its performance metrics using measures such as sensitivity, specificity, and the negative and positive diagnostic likelihood ratios (NLR, PLR). PLR is defined as sensitivity / (1 – specificity), the ratio of the true positive

[4] Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 1999; **8** (2): 135–60.

[5] Pepe M, *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford, 2003; Zhou XH, Obuchowski NA, and McClish DK, Statistical Methods in Diagnostic Medicine Bias, 2nd ed, Wiley, 2011; Zou K, *Statistical Evaluation of Diagnostic Performance – Topics in ROC Analysis*, Chapman & Hall/CRC, 2012; Krzanowski WJ, and Hand DJ. *ROC Curves for Continuous Data*, Chapman & Hall/CRC, 2009; Zweig MH, and Campbell G. Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine, *Clin Chem* 1993; 39(4): 561-577.

[6] Obuchowski NA. An ROC-type measure of diagnostic accuracy when the gold standard is continuous-scale. *Stat Med* 2006; 25: 481-493; Obuchowski NA. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary, *Acad Radiol* 2005; 12:1198–1204.

fraction to the false positive fraction. NLR is defined as (1 – sensitivity) / specificity, the ratio of the false negative fraction to the true negative fraction. Larger values of PLR and smaller values of NLR indicate better classification of the condition status. PLR and NLR are proportional to the odds that test positive and negative subjects have a condition, respectively.

Corresponding two-sided 95% confidence intervals should be provided. The method used to estimate these measures and their corresponding 95% confidence intervals should be clearly pre-specified. If multiple measurements are obtained per subject, the statistical analysis (e.g., 95% confidence interval) should account for the correlation structure of the within-subject measurements using a valid statistical method.

3. Analytical imprecision for qualitative devices

Imprecision is a quantitative value that indicates the extent of disagreement or variability of a set of replicate measurements. For radiation biodosimetry devices providing a quantitative or continuous value (e.g., radiation dose or level), the standard deviation and coefficient of variation are common measures of imprecision. Repeatability is the imprecision when repeated measurements are taken under the same conditions of measurement. Intermediate imprecision is the imprecision when the repeated measurements are taken with some conditions intentionally varied (e.g., run, day, operator, instrument, reagent lot). See CLSI document EP05-A3, "Evaluation of Precision of Quantitative Measurement Procedures—Third Edition."

For biodosimetry devices intended to report qualitative results, the percent agreement of the replicate device results with the qualitative result that is expected for the subject or specimen may be reported. Alternatively, a pure measure of imprecision analogous to the standard deviation for continuous replicate results is the Gini index (or Gini variability). The Gini index is the probability that two categorical results in replicate testing (for the same sample) fall into different categories. For J categories, $g = 1 - \sum_{j=1}^{J} p_j^2$, where p denotes the probability that a replicate gives a result from category J. For two categories, $g=1-p^2-(1-p)^2 =2p(1-p)=2Var(X)$ for Bernoulli random variable X that takes values 1 and 0 with probabilities p and $1 - p$. Further details on the Gini index are available in literature references.[7]

4. Missing data

The SAP should describe how missing data will be handled and documented. Reported results can be misleading if subjects with missing measurement results are excluded from the

[7] Agresti (2002), Categorical Data Analysis, 2nd Ed., p. 68, Light and Margolin (1971), An Analysis of Variance for Categorical Data, Journal of the American Statistical Association, 66 (335), pp. 534-544, Bishop, Fienberg and Holland (1975), Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge, MA. Light RJ, and Margolin BH, An analysis of variance for categorical data, *J Amer Stat Assoc* 1971; 66(335): 534-544.

analysis, the report, or both the analysis and the report. All subjects for whom a measurement was attempted need to be accounted for when reporting results. It is important to analyze the impact of missing data on the conclusions obtained from the study. In some cases it may be necessary to assess if study conclusions are robust given the missing data, and in such cases an intent-to-diagnose or ITD analysis can be performed. An ITD analysis includes every subject or specimen, regardless of whether the subject or specimen is missing the radiation biodosimetry device result, the actual dose, the clinical reference diagnosis, or other results from comparators.

F. Feature Selection During Algorithm Development

If your radiation biodosimetry device incorporates multiple pieces of information into an algorithm in order to produce a single output, validation of this algorithm will be important to understand the performance capabilities of the assay. During algorithm development, it is generally important to obtain a trustworthy estimate of the algorithm's performance before the pivotal performance validation study. Cross-validation is a procedure for estimating the performance of an algorithm on the same dataset on which it was developed. The developmental dataset is split repeatedly into training and test datasets, with the algorithm developed on the training set and evaluated on the test set. The performance estimates obtained for the many splits are then averaged. In bootstrap cross-validation, a training set of the same size as the original dataset is obtained by sampling the subjects or specimens with replacement and evaluated on the remaining unselected subjects or specimens.[8] Cross-validation requires that algorithm development be automated, so may not be possible if the algorithm development process has subjective aspects. All steps of the algorithm construction process, including and especially the step of selecting the features (analytes, measurands, etc.) to be used by the algorithm, should be cross-validated, otherwise the performance estimate will likely be biased.[9] Please note that internal cross-validation is not a substitute for pivotal validation in a dataset that is independent of (external to) the datasets used for development. Further, for the pivotal validation, the final version of the test should be used.

G. Electronic Data

You are encouraged to provide an electronic version of the line data with your submission in an appropriate format such that the datasets are well-described and interpretable. These and the associated programs used to generate your results should be included in a format which can be easily transferred into statistical software. The information at the following URL may be helpful as you prepare these materials:

[8] Efron B, and Tibshirani R. Improvements on Cross-Validation: The .632+ Bootstrap Method. *J Amer Statist Assoc* 1997; 92(438): 548-560.

[9] Simon R et al, Pitfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification, *J National Cancer Institute* 2003; 95(1):14-18.

http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/ucm136377.htm

VI. Appendix B: References

1. United States Department of Agriculture (2010). CFR Title 9, Part 2 Animal Welfare. Office of the Federal Register National Archives and Records Administration. Retrieved from:
   - Definitions: http://www.access.gpo.gov/nara/cfr/waisidx_03/9cfr1_03.html
   - Regulations: http://www.access.gpo.gov/nara/cfr/waisidx_03/9cfr2_03.html
   - Standards: http://www.access.gpo.gov/nara/cfr/waisidx_03/9cfr3_03.html
   - Applicable Policies: http://www.aphis.usda.gov/animal_welfare/policy.shtml

2. United States Public Health Service (2002). Policy on Humane Care and Use of Laboratory Animals. Office of Laboratory Animal Welfare, National Institutes of Health. Bethesda, MD. Retrieved from:
   - http://grants.nih.gov/grants/olaw/references/PHSPolicyLabAnimals.pdf

3. United States Government Health Research Extension Act of 1985, US Government Principles for the Utilization of and Care of Vertebrate Animals Used in Testing, Research, and Training. Retrieved from:
   - http://grants.nih.gov/grants/olaw/references/phspol.htm#USGovPrinciples

4. National Research Council Guide for the Care and Use of Laboratory Animals. (1996). Institute of Laboratory Animal Resources Commission on Life Sciences, Washington, D.C. National Academies of Science Press. Retrieved from:
   - http://www.nap.edu/openbook.php?record_id=5140&page=8