Contains Nonbinding Recommendations

Radiation Biodosimetry Medical Countermeasure Devices

Guidance for Industry and Food and Drug Administration Staff

Document issued on April 18, 2016.

The draft of this document was issued on December 30, 2014.

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

Public Comment

You may submit electronic comments and suggestions at any time for Agency consideration to http://www.regulations.gov. Submit written comments to the Division of Dockets Management, Food and Drug Administration, 5630 Fishers Lane, Room 1061, (HFA-305), Rockville, MD 20852. Identify all comments with the docket number FDA-2014-D-2065. Comments may not be acted upon by the Agency until the document is next revised or updated.

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.

GRZAN

Table of Contents

I. Introduction 1
II. Background 2
III. Scope 4
IV. General Considerations for Submission Contents 4
A. Benefit-Risk Analysis 4
B. Device Description and Specifying the Intended Use 5
1. Stage of Response 6
2. Appropriate timeframes for testing 6
3. Assay limitations 6
C. Samples 6
1. Sample availability 6
2. Specimen collection and handling 7
3. Unirradiated samples 8
D. Analytical Validation Studies 8
1. Accuracy in relation to the dose delivered 9
2. Measuring interval or expected values 9
3. Interference 10
4. Other analytical testing protocols 10
E. Controls and Calibrators 11
F. Instrumentation and Software 11
G. Animal Study Considerations 12
1. Defining an appropriate animal model 12
2. Effect of confounding factors on animal studies 13
3. Bridging animal data to human data 14
H. Establishing Performance Characteristics – Clinical Validation Studies with Human Samples 15
1. Prospective clinical studies 15
2. Clinical studies using banked samples 16
3. Limitations of clinical studies 16
I. Labeling 16
1. Instructions for use 17

2. Limitations 17
3. Interpretation of results 18
4. Performance characteristics 18
J. CLIA Categorization 18
V. Appendix A: Statistical Considerations for Biodosimetry Devices 20
A. Study Blinding (Masking) 20
B. Confidence Intervals for Estimates of Performance 20
C. Study Analysis 20
1. Quantitative, continuous, or semi-quantitative output 22
2. Qualitative output 22
3. Analytical imprecision 23
4. Missing data 23
D. Feature Selection During Algorithm Development 23
E. Electronic Data 24

Radiation Biodosimetry Medical Countermeasure Devices

Guidance for Industry and Food and Drug Administration Staff

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff or Office responsible for this guidance as listed on the title page.

I. Introduction

FDA has developed this guidance to provide industry and Agency staff with recommendations for the types of information that should be submitted to support marketing authorization (e.g., clearance or approval) for radiation biodosimetry medical countermeasure devices (referred to as "biodosimetry devices" or "biodosimeters" throughout this document). This guidance applies to premarket submissions for medical device systems intended to measure biological responses to unintended (non-therapeutic) radiation absorption. Biodosimetry 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. Biodosimetry technologies may be used at various stages during triage, including both early mass casualty triage and subsequent clinical evaluation. Such exposures could be the result of intentional harm or as a 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 guidance document.

This 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 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. Many of the considerations included in this guidance document assume that the biodosimeter output is related to a quantitative or qualitative measure of dose. We acknowledge that there may be other relevant device outputs that manufacturers would like to pursue, such as risk of developing specific clinical symptoms (e.g., related to acute radiation syndrome). Many of the items discussed below can be applicable to a diversity of device outputs, if suitably addressed in a

premarket submission. However, device outputs related to, or predicative of, clinical symptoms or outcome may be challenging to evaluate in the premarket setting because of the paucity of human data and challenge of translating results from animal studies to human use scenarios. In this case, we would recommend early engagement with FDA via the pre-submission process. Additionally, because it may be impossible to validate biodosimetry devices in the intended use population, sponsors should plan to include in their ultimate pre-market submission a plan or approach to conduct post-market testing when an emergency arises and their product is utilized in a real-world setting. 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 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 as a consequence of a natural disaster that causes a large-scale industrial radiological accident, could potentially expose thousands of individuals to high levels of radiation that would require immediate assessment and medical intervention. A coordinated medical response including both early mass casualty triage and subsequent clinical evaluation would be important to mitigate harm resulting from the non-therapeutic exposure of individuals to ionizing radiation.[1] Biodosimetry tools would be a critical component of such a response and will serve to differentiate the acutely irradiated population from the concerned public. Radiation biodosimetry estimates the absolute dose of radiation absorbed by an individual. Biodosimetry devices provide information that allows for the development of appropriate treatment plans for patients.

The most significant benefit that biodosimetry has over physical dosimetry is that it takes into account the natural patient biological variability in radiation response, though physical dosimetry tools may provide a more accurate measurement of the actual radiation dose delivered. For example, a physical dosimeter such as a thermoluminescent dosimeter detects radiation on a physical substrate but does not assess the amount of radiation absorbed by the subject or the subject's biological response to radiation. Therefore, while a physical dosimeter may provide an accurate representation of dose, these devices will be unable to differentiate a patient who is radiation-sensitive (presumably needing a higher level of medical intervention) from a patient

[1] For more information on the Federal government response plans for radiation disaster scenarios and a list of the applicable drugs and treatments recommended in such a scenario, see the document entitled "Planning Guidance for Response to a Nuclear Detonation" (http://www.remm.nlm.gov/PlanningGuidanceNuclearDetonation.pdf). This document was developed by the National Security Staff Interagency Policy Coordination Subcommittee for Preparedness and Response to Radiological and Nuclear Threats (last accessed on April 14, 2016). Note this website is not controlled by FDA.

who is radiation-resistant. As such, the development of radiation biodosimetry techniques will be a powerful tool in personalizing therapeutic responses to radiation absorption, representing a significant benefit to public health.

Most biodosimeters are *in vitro* diagnostic devices (IVDs), as defined in 21 CFR 809.3(a).[2] A wide array of technologies may be employed to assess biological responses to radiation. Methodologies amenable to incorporation into 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 other methodologies designed to detect other biological signals induced by absorption of radiation.

Sponsors who intend to market 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 marketing authorization (e.g., premarket clearance or approval) prior to marketing their devices. This guidance document represents our current thinking regarding the recommended information that should be included in a premarket submission of a biodosimetry device for assessing unintended absorption of 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). The following documents may be useful in preparing premarket submissions:

- "Format for Traditional and Abbreviated 510(k)s" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm084365.htm)
- "Acceptance and Filing Reviews for Premarket Approval Applications (PMAs)" (http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM313368.pdf)
- "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)

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

[2] *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.

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 biodosimetry devices. The intent of this guidance document is to provide clarity to both industry and FDA staff on performance data and information that should be included in submissions to the Agency 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 biodosimetry device development. The pre-submission process is an opportunity for sponsors to get regulatory advice on their product development plan. The pre-submission process, in general, leads to better prepared submissions and shorter review times. 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 guidance only applies to validation of diagnostic biodosimetry devices intended to be used to assess radiation absorption that occurs as a result of non-therapeutic or accidental exposures (e.g., a deliberate attack, such as use of an improvised nuclear device, or a natural disaster), and does not apply to medical devices intended to be used to measure doses delivered as a result of radiation therapy nor to devices that measure effects from long-term radiation 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. General Considerations for Submission Contents

A. Benefit-Risk Analysis

A source of significant risk to patient health associated with biodosimetry devices is when 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 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 absorption in an emergency scenario could lead to unnecessary or inappropriate treatment for acute radiation syndrome[3] or may result

[3] Acute Radiation Syndrome (ARS), also referred to as radiation sickness, is an acute set of illnesses caused by irradiation of the entire body (or most of the body) by a high dose of penetrating radiation in a very short period of time (usually a matter of minutes). According to Centers for Disease Control and Prevention (CDC), ARS is

in the patient being inappropriately placed in an expectant category and given only palliative care when treatment could be life-saving. However, a false negative result or underestimation of absorption could lead to failure to provide treatment or incorrect patient management which may be lethal. Thus it is essential to appropriately balance the benefits of biodosimetry devices with the risks, including the risk of incorrect 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 biodosimetry devices includes consideration of the overall time to result or throughput of the system (e.g., decreased accuracy within a reasonable range 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. This discussion should include but not be limited to 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 biodosimetry 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), 3) the specific population(s) for which the test is intended (e.g., pediatrics, general population), and 4) the intended use setting(s) (e.g., field triage or professional use only). For the purposes of this guidance, field triage devices would be intended to be used by trained lay personnel (first responders) in a setting outside of a professional medical facility.

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 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 and included as part of the intended use statement or in the device description for biodosimetry devices include the following:

---

classified into three classic sub-syndromes: hematopoietic syndrome (H-ARS), gastrointestinal syndrome (GI syndrome), and cardiovascular/central nervous system syndrome (CV/CNS syndrome) (http://www.bt.cdc.gov/radiation/arsphysicianfactsheet.asp). Note this website is not controlled by FDA.

1. Stage of Response

Biodosimetry devices may be designed for early field triage (usually qualitative). For these types of devices, the intended use should specify the throughput capabilities, time to result of the device, and the decision making cut-points assessed. Alternatively, biodosimetry devices may be designed for subsequent clinical evaluation for more definitive medical management, including dose level confirmation (usually quantitative or semi-quantitative). For these types of devices, specific clinical indicators of health status should be part of the intended use statement and, for quantitative devices, should include the assay measuring interval.

2. Appropriate timeframes for testing

Many common biomarkers of radiation absorption display defined kinetics, during which they become detectable, remain stable, and then disappear from the matrix being examined. Therefore, it is important to specify the timeframe 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 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. Limitation statements may be included in the intended use statement or in the warnings and contraindications section of the labeling, depending upon the risks associated with using the device in those limited situations. For instance, if validation testing was only performed on certain populations (e.g., not tested in pediatrics), for specific radiation types (e.g., gamma source only or limited to effects from low-linear energy transfer [LET] radiation), 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 for the intended use population. You should make every attempt to establish assay performance in 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. Samples

1. Sample availability

We recognize the difficulty in acquiring appropriate clinical samples for use during device validation. Typically, the source of clinical samples is patient samples taken in the course of radiation therapy (partial or total body). However, these specimens may be difficult to obtain in

sufficient amounts. Another challenge with these sources is that the therapeutic protocol restricts the radiation doses available (only low dose fractionated exposures may be readily obtained) and such specimens may not be sufficient to represent the spectrum of radiation absorption that occurs as a result of non-therapeutic or accidental exposures. In the absence of sufficient or appropriate clinical samples, contrived samples may be used for analytical performance testing, but only as a supplement to, and not as a substitute for clinical samples.

Samples may be contrived by a variety of means including but not limited to:

- *ex vivo* irradiation of the appropriate matrix,
- spiking the analyte of interest into the appropriate matrix,
- through the use of animal-derived specimens (e.g., specimens obtained from irradiated animals), as appropriate.

In some instances, control materials (i.e., materials designed to serve as assay positive or negative controls) or contrived samples (i.e., samples designed to mimic a patient sample) 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. You should make every effort 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, time of collection after exposure, 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 also be addressed in the premarket submission. However, we recognize there may be different analytical and clinical performance needs depending on whether the device is intended for early mass casualty triage in a field setting, for radiation absorption confirmation, or for dose refinement in a clinical laboratory setting.

For example, if the device is intended to be used in an early mass casualty triage environment in a field setting, 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, 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. Unirradiated samples

Unirradiated samples may be used both to assess the normal range (reference interval) or expected values for qualitative or semi-quantitative assays. In addition, these samples may be used to create contrived samples by spiking the analyte of interest into an appropriate matrix. Assessment of the reference interval or expected values of analyte expression should include a suitable number of normal samples to determine if biodosimeter analyte expression depends on subject age, race, sex, and common disease conditions (e.g., obesity, diabetes, or arthritis, as appropriate). Normal samples can be obtained prospectively or from appropriate sample banks. All normal samples should be collected as intended for your device. For instance, if blood will be collected in ethylenediaminetetraacetic acid (EDTA) tubes for use in your biodosimeter, then normal samples should be collected in EDTA tubes. As above, basic demographic information and information on disease conditions should be collected with normal samples.

D. Analytical Validation Studies

Biological response to a radiation dose as measured by biodosimetry is confounded by patient-to-patient variation in radiation response, 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 of the assay output to the radiation dose delivered 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 measuring interval of the device in a defined sample subset. This analytical performance information will be critical to substantiate intended use claims of a biodosimetry device.

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 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. Accuracy in relation to the dose delivered

In order to demonstrate analytical accuracy, if the device output is radiation dose, then the device's measurement of the biological response to radiation will probably be 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 be confounded by the biological response to radiation, there will be between-individual variations which may complicate the correlation of the output to the radiation dose delivered.

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 clinical decision making cut-points relevant to the intended use of your device (i.e., mass casualty triage or subsequent clinical triage for more definitive medical management) and contain pre-specified success criteria. The rationale for your acceptance criteria should be provided and clinically justified.

A manufacturer may consider using a reference method to validate the accuracy of their biodosimetry device. For example, the dicentric chromosome assay (DCA) may be an appropriate reference method in some circumstances. However manufacturers should understand the limitations of available reference methods. For instance, DCA has several limitations, including that above 5 Gy the assay is no longer dose-dependent. Because of these limitations, sponsors may also choose not to assess assay performance compared to a reference method.

2. Measuring interval or expected values

a) Measuring interval

If a biodosimetry device reports numerical (quantitative) results, submissions should include studies establishing the measuring interval of your device. Measuring interval studies should be designed to substantiate Limit of Detection (LoD), Limit of Blank (LoB), Limit of Quantitation (LoQ), and the linearity of the assay. The radiation absorption levels that are clinically relevant span from 2-10 Gy, and we recommend developing test protocols to evaluate the limitations of your assay outside (both above and below) the clinically relevant dose range. For instance, if high radiation levels can cause the assay to incorrectly report lower values (i.e., the hook effect), then the limitations section of the labeling should include this information. Please refer to the CLSI document EP6-A and EP17-A2 to assist in the design of such studies.

b) Category assessments

If a biodosimetry device reports a qualitative or semi-quantitative result, submissions should include studies validating the decision making cut-point(s) and how outputs will be assigned to clinical categories. Samples around critical decision making cut-points should be included in these studies.

3. Interference

Interfering substances may affect assay outputs. As such, biodosimetry device submissions should include the results of appropriate interference studies (see CLSI Document EP07-A2) for quantitative tests. Please include samples near the cut-off(s) for interference testing in qualitative or semi-quantitative devices. 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. For more information on the Federal government response plans for radiation disaster scenarios and a list of the applicable drugs and treatments that will be recommended in such a scenario, see the document entitled "Planning Guidance for Response to a Nuclear Detonation."[4]

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 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.G.2 below).

4. Other analytical testing protocols

Analytical testing protocols, in addition to those mentioned specifically above, should be included in biodosimetry submissions as applicable given the intended use, output, and technology. Reproducibility should generally be demonstrated at a minimum of three sites (of which at least one should be in the United States) and submissions should include an evaluation of sources of variability including operator-to-operator, instrument-to-instrument, day-to-day,

[4] This document was developed by the National Security Staff Interagency Policy Coordination Subcommittee for Preparedness and Response to Radiological and Nuclear Threats (last accessed on April 14, 2016). Note this website is not controlled by FDA.

and kit lot-to-kit lot variability, as applicable. Repeatability should also be evaluated. The LoD (limit of detection), LoQ (limit of quantification), and LoB (limit of blank), as relevant, may be demonstrated in a separate study for quantitative devices. The stability of kit reagents, including controls and calibrators, as applicable 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:

- EP05-A3, "Evaluation of Precision of Quantitative Measurement Procedures"
- EP09-A3, "Measurement Procedure Comparison and Bias Estimation Using Patient Samples"
- EP12-A2, "User Protocol for Evaluation of Qualitative Test Performance"
- EP25-A, "Evaluation of Stability of *In Vitro* Diagnostic Reagents"

E. Controls and Calibrators

The design of 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).

F. Instrumentation and Software

If your 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/downloads/MedicalDevices/.../ucm073779.pdf). 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 referenced above.

G. Animal Study Considerations

Typically, 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 clinical trial or to use properly archived surplus (excess) specimens to demonstrate test performance when prospective studies are not feasible. However, in the case of 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:[5]

- 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 study is either unethical, or prospective studies that may be ethically performed will not yield a sample set adequate to demonstrate assay performance over the measuring interval 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 adequate substitute for human studies. In particular, establishing that there is high homology in the analyte(s) being assessed and that the animal model displays similar responses to radiation in the biological pathways being interrogated is important. We recommend that you provide a justification (e.g. appropriate published literature) that the radiation dose given in your selected animal model is comparable to the dose in humans.

When choosing your animal model, you should consider transferability of experimental results, the influences of genetic uniformity of the model, and background knowledge of radiation response pathways. For example, while rodent models may be appropriate for early proof of concept studies in biodosimetry, their biological divergence from humans presents challenges for demonstrating biodosimetry effectiveness in humans. Large-animal models (non-human primate, canine, or porcine), in which radiation response more closely resemble those of humans, may be better options for effectiveness studies. You should ensure that an equal distribution of sexes is included in pre-clinical testing and consider including animals at various age ranges so age and sex-related differences in radiation response can be assessed for the analytes included in the biodosimeter. You should carefully determine the number of animals that are required for your

[5] Please note that this guidance refers to the use of animal model data for biodosimetry device submissions only and is not applicable to the *Animal Rule* pathway for approval of drugs (21 CFR Subpart I) or biological products (21 CFR Subpart H). For information and recommendations on drug and biological product development under the *Animal Rule*, see FDA's final guidance *Product Development Under the Animal Rule*, Guidance for Industry, available at http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm399217.pdf.

studies so that meaningful data are collected. Consider consulting a biostatistician for assistance with determining the required animal numbers for use in these studies.

Multiple animal models generally would not be required to demonstrate effectiveness. However, multiple animal models may be necessary when a single appropriate animal model cannot be identified for all the analytes being assessed by the device. Developers can perform a literature search in order to determine the most appropriate animal models to be used in biodosimetry device effectiveness studies.

Please note that all device studies using animal models that are intended to support premarket submissions to the Agency must comply with all applicable laws and regulations including:

- The Animal Welfare Act (7 U.S.C. 2131 et seq.) and related regulations (9 CFR Part 2)
- The Public Health Service Policy on Humane Care and Use of Laboratory Animals[6]
- Good Laboratory Practice for Nonclinical Laboratory Studies regulations (21 CFR Part 58, including 21 CFR 58.90)[7]
- U.S. Government Principles for the Utilization of and Care of Vertebrate Animals Used in Testing, Research, and Training[8]
- Guide for the Care and Use of Laboratory Animals[9]

Further, IVD sponsors should follow the guidelines below:

- AVMA Guidelines for the Euthanasia of Animals[10]
- Recognition and Alleviation of Pain in Laboratory Animals[11]

2. Effect of confounding factors on animal studies

In some cases, animal housing environmental conditions, veterinary care, and supportive care might influence biodosimetry assay results. For instance, analyte expression might be influenced

[6] U.S. Department of Health and Human Services, National Institutes of Health, Office of Laboratory Animal Welfare, 2015, available at http://grants.nih.gov/grants/olaw/references/PHSPolicyLabAnimals.pdf (accessed on April 14, 2016).

[7] While the GLP regulations do not specifically address data quality and integrity for animal studies for the accuracy of IVDs, FDA considers them to be a well-established and relevant system for ensuring data quality and integrity for animal efficacy studies. FDA recommends the use of the GLP regulations to ensure the quality and integrity of data from these studies.

[8] U.S. Department of Health and Human Services, National Institutes of Health, Office of Laboratory Animal Welfare – Public Health Service Policy on Humane Care and Use of Laboratory Animals, 2015, available at http://grants.nih.gov/grants/OLAW/references/phspol.htm (accessed on April 14, 2016).

[9] National Research Council of the National Academies, Committee for the Update of the Guide for the Care and Use of Laboratory Animals, Institute for Laboratory Animal Research, 2011, available at http://grants.nih.gov/grants/olaw/Guide-for-the-Care-and-Use-of-Laboratory-Animals.pdf (accessed on April 14, 2016).

[10] American Veterinary Medical Association, *AVMA Guidelines for the Euthanasia of Animals: 2013 Edition*, 2013, https://www.avma.org/KB/Policies/Documents/euthanasia.pdf (accessed on April 14, 2016).

[11] National Research Council (US) Committee on Recognition and Alleviation of Pain in Laboratory Animals, *Recognition and Alleviation of Pain in Laboratory Animals*, 2009, Washington, DC: National Academies Press.

by differences in diet or housing conditions of laboratory animals as opposed to that of the general human population. You should standardize husbandry of animals used in these studies, offering the same care, diet, and standardized enrichment to animals included in the study in order to help minimize experimental inconsistencies. Additionally, to avoid variability caused by changes in radiation response due to circadian rhythm patterns, we recommend you deliver radiation at the same time of day (e.g., a.m. vs. p.m.) to all animal models.

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. For example, you could consider providing supplemental care such as parenteral and/or oral physiologic fluids, as well as soft, nutrient-rich palatable foods that encourage eating. You may also consider providing supplemental heat, as well as antibiotics when indicated to determine if these types of medical interventions alter the biological responses being assessed by biodosimetry.[1] For example, if neutrophil numbers are part of a biodosimetry algorithm, it should be understood how G-CSF (granulocyte-colony stimulating factor) treatments affect resulting radiation 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 biodosimeter at a clinically relevant dose or dose range, 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 studies 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. There are several complex statistical considerations in determining the species comparison that will be appropriate depending on device output (e.g., quantitative, qualitative, or semi-quantitative) and technology type. These statistical methodologies should be discussed with the Agency on a case-by-case basis. 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. Depending on the mechanism of action or biological pathways assessed by the device, you could consider testing animal models or providing literature support for the impact that commonly used drugs such as anti-cholesterol, anti-hypertension, diabetic drugs, and other common drugs may have on the analyte (e.g., a particular metabolic pathway, CYP450, or growth factors) being evaluated. Other situations that may confound assay results such as combined injury can also be examined in animal models.

Animal studies should be designed to use the fewest possible animals to demonstrate statistical significance in accordance with Institutional Animal Care and Use Committee (IACUC) guidelines.[12] Because these studies will be critical to demonstrate the performance of a

[12] *See* note 8.

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 biodosimetry study designs.

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

FDA expects that biodosimeter safety and effectiveness will be established in an appropriately designed clinical study that captures the device performance in human subjects. However as discussed above, non-clinical models may be used to supplement human clinical data by providing information on doses, dose rates, time-courses and radiation sources that cannot be ethically obtained in a human clinical study. 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

Studies may be performed to prospectively assess biodosimeter performance in the context of radiation delivered for therapeutic purposes. Because patient numbers may be limited, you should consider a study design that collects multiple samples from each patient over the course of their therapy. Additionally, whenever possible you should collect specimens prior to radiation and/or other therapy as a clinical baseline. Ideally, you should design studies to assess biodosimeter performance beyond the range of times post-exposure that you intend to claim in labeling. For example, if you intend your device to have labeling claims from 24 hours to 7 days post-exposure, your studies should include a minimum of 8 days post-exposure. However, we acknowledge that therapies administered pre-, post-, or concurrently with radiation in some patient populations may confound biodosimetry test results. As such, lack of human clinical data across the intended times post-exposure may be acceptable with an appropriate scientific justification and accompanying supplemental animal data.

For all prospective clinical studies, information should be captured on the dose, dose rate, time-course and radiation source for all patients. In addition, basic demographic information should be collected, including patient age, sex, and race. Information on the patient's clinical condition (e.g., body mass index and medications) should also be collected whenever possible.

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).

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. Clinical studies using banked samples

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 samples (collected for other purposes) 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. Limitations of clinical studies

We acknowledge that there may be significant limitations in the interpretation of clinical studies for 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.

I. Labeling

The labeling of a 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"

- In Vitro Diagnostic Device Labeling Requirements (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 biodosimetry devices that complies with the labeling requirements outlined in 21 CFR Parts 801 (for general labeling requirements) and 809 (for additional IVD labeling requirements), as applicable.

1. Instructions for use

For 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 using neutron radiation sources, then this information should be included in the labeling.

3. Interpretation of results

Labeling must include a description of the expected values for your device (qualitative) or measuring interval (quantitative) (see 21 CFR 809.10(b)(11)). In the case of a qualitative device, we recommend that the expected values be portrayed in terms of expected values for non-irradiated subjects, and around the clinical decision making cut-points that were evaluated for your device (e.g., 2 Gy and/or 10 Gy). If the results are qualitative, you should explain how to interpret positive and negative results, including their clinical significance. Alternatively, if the device is quantitative, you should explain the clinical significance of the outputs along the measuring interval. 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.

J. CLIA Categorization

The CLIA (42 U.S.C. 263a) regulates laboratory testing and requires that clinical laboratories obtain a certificate before accepting materials derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease or the impairment of, or assessment of the health of human beings. The type of CLIA certificate a laboratory obtains depends upon the complexity of the tests it performs. CLIA regulations describe the following three levels of test complexity: waived tests, moderate complexity tests, and high complexity tests, 42 CFR 493.5(a).

CDRH determines the CLIA categorization of IVD tests at the time of review of a premarket notification submission (510(k)) or a premarket approval application (PMA). Biodosimeters intended for field triage that are cleared for home use for the purposes of CLIA are CLIA waived (42 CFR 493.15). In addition, a biodosimeter meets the statutory criteria for CLIA waiver if the sponsor demonstrates in studies conducted at the intended use site that the biodosimeter employs methodologies that are so simple and accurate as to render the likelihood of erroneous results negligible or that the biodosimeter poses no reasonable risk of harm to the patient if the test is performed incorrectly. If a sponsor of a test categorized as moderate complexity believes their

test meets the statutory criteria for CLIA waiver, they may submit a CLIA Waiver by Application to request categorization of the test system as waived. A Dual 510(k) and CLIA Waiver by Application pathway is also available. For additional information regarding CLIA Categorizations and CLIA Waivers by Application including the Dual 510(k) and CLIA Waiver by Application pathway, refer to the guidance entitled "Administrative Procedures for CLIA Categorization" (http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm070762.htm).

V. Appendix A: Statistical Considerations for Biodosimetry Devices

This appendix includes some statistical considerations for biodosimetry devices. Further statistical considerations for diagnostic devices that may be applicable to 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. Study Blinding (Masking)

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

B. Confidence Intervals for Estimates of Performance

Confidence intervals for performance estimates are also referred to as precision of estimation in clinical study literature. Performance variability 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.

C. Study Analysis

The protocol for a 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 (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.

In clinical studies of IVDs, laboratories often produce results that are not described by a defined category or numerical value. The terms used to describe these results include "uninterpretable," "invalid," and "indeterminate." Your test results may also be called equivocal. For various reasons including inability to obtain new samples after assay failure, results may be missing. The SAP should include a plan for dealing with all these possible device results 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.C.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 (e.g., for futility, 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.[13]

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.

[13] 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.

1. Quantitative, continuous, or semi-quantitative output

Statistical analyses and methods depend on the type of results provided by the biodosimetry device. For devices providing a quantitative measurement, the bias and imprecision of the measurement should be evaluated along with other performance characteristics. For example, difference plot methodology can be used to compare the level of radiation absorption predicted by the device with the actual level delivered (actual dose or reference level) over the measuring interval.[14] Comparison of device output and delivered radiation dose results near clinical decision making cut-points is essential. The systematic difference (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 EP05-A3, "Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline--Third Edition."

Receiver Operating Characteristic (ROC) analysis might also be useful to evaluate the diagnostic accuracy of a 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 cut-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.[15] If the condition of interest is not binary (e.g., dose or level of radiation absorption), generalizations of AUC exist for evaluating the discrimination ability of medical tests.[16]

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

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

[15] 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.

[16] 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.

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 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.

3. Analytical imprecision

Imprecision describes the extent of disagreement or variability of a set of replicate measurements. For 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.

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 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.

D. Feature Selection During Algorithm Development

If your 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.[17]

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.[18] 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.

E. 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:

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

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

[18] 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.