FDA Advisory Committee Briefing Document

Cardiovascular and Renal Drugs Advisory Committee

December 10, 2008

New Drug Application 22-349
Imagify™ (perflubutane polymer microspheres)

Table of Contents

	Page
1. Introduction	2
2. Topics for Questions and Discussion	4
3. Data Summary	7
4. FDA Guidance Document Bibliography: Developing Medical Imaging Drug and Biological Products:	21
Part 1. Conducting Safety Assessments	
Part 2. Clinical Indications	
Part 3. Design, Analysis, and Interpretation of Clinical Studies	

1. Introduction

Ultrasound contrast agents are "microbubble" or "microsphere" drugs that are usually injected into the vascular system in order to improve the diagnostic information obtained from ultrasound-based images. The gas contained within these products enhances the acoustic signal and provides contrast between the vascular system and surrounding structures.

"Microbubble" drugs consist of a gas core encased within a molecular shell, such that the bubbles oscillate in response to ultrasound waves. AI-700 is a "microsphere" drug, in that it consists of a mechanically "rigid" particle which contains pores that encase a gas. The relative inflexibility of the microsphere is proposed to allow longer ultrasound visualization, in comparison to microbubble products. Because the AI-700 microspheres are generally smaller than a red blood cell, injection of AI-700 into a vein results in the passage of the product through the pulmonary vasculature and into the arterial system, including the left ventricle and coronary arterial system. Dispensation of AI-700 within the arterial system allows ultrasound visualization of the left ventricular cavity and perfused myocardial tissue.

AI-700 (Imagify™) is the subject of a New Drug Application (NDA) with the following proposed indication:

"Imagify is an ultrasound imaging agent indicated for patients with stable chest pain being evaluated for inducible ischemia for the detection of coronary artery disease (CAD) based on assessment of myocardial perfusion and wall motion. Imagify echocardiography is accomplished with rest and stress techniques."

In general, the NDA sponsor has proposed that AI-700-echocardiography can be used as a screening tool to assist the clinician in the stratification of patients for referral to coronary arteriography. At the present time, alternative modalities that serve a similar purpose include exercise stress testing with electrocardiographic monitoring, echocardiographic stress testing (non-contrasted) as well as stress testing with radionuclide myocardial perfusion imaging, such as the performance of single photon emission-computerized-tomography (SPECT) following administration of Tc99m myocardial imaging agent. The stress testing may be performed using exercise or administration of a drug to "stress" the cardiovascular system (pharmacologic stress, using dipyridamole, adenosine or dobutamine).

As described in the attached guidance documents, FDA's determination of the diagnostic efficacy of an imaging agent is based upon the agent's performance characteristics (sensitivity and specificity with respect to a "standard of truth" *or* extent of agreement with a reference test), the proposed claim (or indication) and an overall assessment of the agent's risks with respect to the diagnostic benefit. The confirmatory clinical studies for AI-700 assessed the performance characteristics (sensitivity, specificity, accuracy) with respect to a standard of truth defined by coronary arteriography or clinical history (such as documented history of myocardial infarction).

In the main AI-700 clinical studies, a single arm design was used in which patients underwent rest and stress echocardiography as well as rest and stress SPECT following Tc99m administration. Dipyridamole was used to induce pharmacologic stress. In general, the primary

endpoints related to verification that AI-700 performance characteristics were not importantly inferior to the SPECT results. With respect to efficacy considerations, this approach is consistent with FDA guidance document recommendations. However, one of the challenges of this approach is the lack of a comparator group for safety evaluations. An additional challenge for evaluating the AI-700 studies relates to the confounding of potential drug-related adverse events with adverse events related to the pharmacologic stress agent.

FDA is particularly concerned about the safety of AI-700 because of signals evidenced in animal and clinical studies. These concerns appear in the setting of accumulating postmarketing data regarding the risks associated with the use of microbubble ultrasound contrast agent. FDA concerns regarding the safety of microbubble ultrasound contrast agents were the subject of a June, 2008 Advisory Committee discussion where the FDA provided a summary of uncommon but serious or lethal hypotensive reactions observed in the post-marketing experience for the microbubble products. That committee's concerns focused upon the need for comparative clinical studies to facilitate the interpretation of the adverse events that may follow administration of a contrast agent.

The FDA's review of the AI-700 animal and clinical data is ongoing. The main goal of this advisory committee is to obtain the committee's perspective regarding the data, as presented by the company and summarized by the FDA, in the context of FDA's preliminary concerns. The committee's perspectives regarding the more readily obvious concerns will importantly inform the remainder of FDA's review.

To date, FDA is concerned that the available efficacy data do not provide a consistent (between the two main clinical studies) characterization of AI-700 performance characteristics and the available performance characteristics are not sufficient for the proposed diagnostic use of the product. FDA is also concerned that the safety risks associated with the product are inappropriate for the product's proposed clinical use.

2. Topics for Questions and Discussion

FDA anticipates questions and discussions related to the following topics:

1. Two clinical studies were performed to establish the efficacy of AI-700 as an ultrasound contrast agent for use in the detection of coronary artery disease, based on assessment of myocardial perfusion and wall motion at rest and following pharmacologic stress. Study 32 enrolled patients with a history of chest pain who had clinical indications for rest and stress SPECT myocardial perfusion imaging. Study 33 enrolled patients with a history of chest pain who were scheduled to undergo coronary arteriography or had already undergone coronary arteriography within the past 30 days. In both studies, AI-700 results were compared to the results obtained with SPECT, using a truth standard (angiography and/or clinical criteria). AI-700 performance characteristics differed between the two studies and neither study evaluated the "added value" of AI-700 contrast enhanced ultrasound to non-contrasted ultrasound. FDA anticipates asking questions related to:

a. The extent to which the performance characteristics provide clinically important and consistent (between the two studies) evidence of diagnostic efficacy;

b. The extent to which the diagnostic efficacy is consistent with the proposed clinical use of AI-700. Within the NDA documents, the sponsor has proposed that, "AI-700 is intended for use in stable chest pain patients as a screening tool to stratify patients for referral to coronary arteriography."

c. The clinical importance of the data's inability to establish the "added value" of AI-700 ultrasound over non-contrasted ultrasound.

2. AI-700 consists of physically rigid particles, most of approximately 2 microns in size. The gas, perflubutane is contained within the particles and provides the basis for the ultrasound contrast. After injection into the blood stream, elimination of AI-700 appears to use the reticuloendothial system (particularly macrophages) for engulfment/metabolism of the physically rigid portions of the AI-700 particles while the perflubutane portion is exhaled. Administration of AI-700 to healthy volunteers resulted in systemic release of biomarkers associated with inflammation (C3a, C-reactive protein and white blood cell count alterations).

AI-700 was administered to 1,194 subjects, including 911 patients with known or suspected coronary artery disease who received the proposed dose regimen. Most adverse events were reported following stress testing and interpretation of these events is confounded by the stress testing procedure, including administration of dipyridamole.

Among the population of patients with known or suspected coronary artery disease who received the proposed dose regimen:

- No deaths were reported.

- Serious adverse events were reported in 11/911 (1%) of patients, including 3 patients who experienced "vasovagal syncope" shortly following the first injection of AI-700.

- 16/911 (2%) had the AI-700 injection permanently discontinued because of adverse events (e.g., cyanosis, hypotension, rigors). Most of these patients (14) experienced the events during the initial AI-700 injection.

- 7/911 (1%) had the AI-700 injection temporarily interrupted or the dose regimen adjusted because of adverse events (e.g., chest pain, hypotension, bradycardia). All seven of the patients experienced the events during the second AI-700 injection.

- 38/911 (4%) experienced adverse events of "hypotension or decreased blood pressure" with 12 (1%) of the patients experiencing the events in the time period between initiation of the first AI-700 injection and the administration of dipyridamole.

- The most common adverse events (≥ 1%) in the time period between initiation of the first AI-700 injection and the initiation of dipyridamole were: headache (3%), flushing (2%), rigors (1%) and hypotension (1%).

- The most common adverse events related to laboratory abnormalities were: leukocytosis (7%), neutrophil count increases (6%) lymphocyte decreases (2%), and monocyte decreases (1%). Most of these adverse events were reported in the time period following administration of the second AI-700 injection to the time of discharge (generally 2 - 3 hours following the second AI-700 injection).

FDA anticipates asking questions related to:

a. Do the biomarker indicators of systemic inflammation in healthy subjects (C3a, C-reactive protein, WBC count alterations) provide evidence of potentially clinically important systemic inflammation? If so, do the available data from patients sufficiently characterize the risks for systemic inflammatory reactions?

b. Is the sample size of the safety database and the patient characteristics sufficient to characterize the major risks associated with AI-700?

c. What clinical importance do you assign to the types of serious and other adverse events reported following AI-700 administration? For example, do you consider the nature and frequency of these events typical of those expected for an imaging agent used as a screening tool to stratify patients for referral to coronary arteriography?

3. Does the diagnostic efficacy and potential clinical usefulness of AI-700 justify exposure to the risks associated with the product?

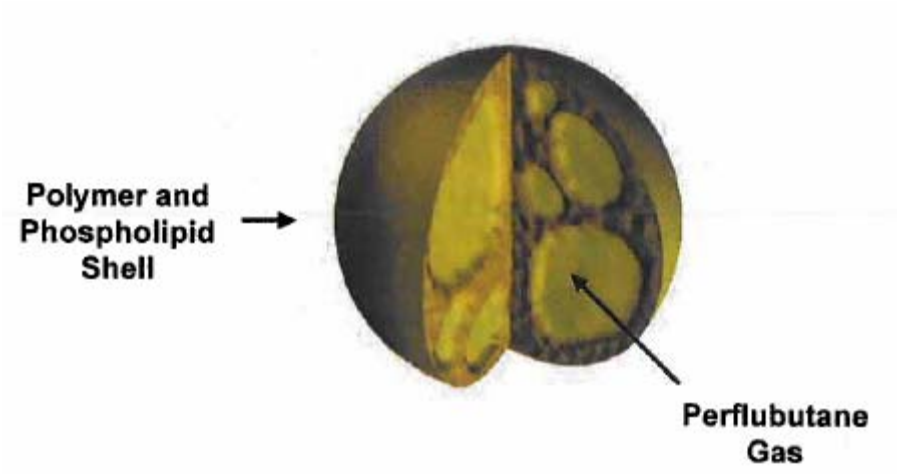
FDA may request discussion of the need for any additional studies, either as pre-marketing or post-marketing expectations.

3. FDA Interim Data Summary

a. Product Characteristics and Initial Development:

AI-700 is composed of the gas, perflubutane (*n*-decafluorobutane) (PFB), contained within porous microspheres (Figure 1). The microspheres are manufactured from the polymer poly (D,L-lactide-co-glycolide) and the phospholipid 1,2-diarachidoyl-*sn*-glycero-3-phosphocholine. The microspheres have a mean diameter of approximately 2 µm; ≥ 99.0% of particles are ≤ 10 µm, and 100.0% are ≤ 20 µm in diameter. AI-700 is supplied as a sterile, lyophilized powder in a 20-ml, single-use glass vial. AI-700 is reconstituted with 5 mL of Water for Injection (WFI), to form an iso-osmotic suspension that is administered as a slow bolus intravenous injection for up to 10 minutes, in duration.

Figure 1 : AI-700 microsphere [drawing provided by Sponsor]



Nonclinical safety testing of AI-700 revealed the following safety concerns when AI-700 was tested:

- in pigs, a "no observable effect" level could not be defined because of acute blood pressure changes (pulmonary artery hypertension and systemic hypotension), including changes that occurred at clinically-relevant AI-700 doses. The hemodynamic changes were brief and generally resolved within 10 to 15 minutes following cessation of dosing. The effects were generally similar to those observed with a microbubble ultrasound contrast agent. The meaningfulness of the pig as a model for cardiovascular effects has been questioned because the animals have more pulmonary intravascular macrophages than most other species.

- in monkeys at doses higher than those to be administered clinically, mean arterial blood pressure decreased 15 to 25% from baseline. Clinically-applicable doses were not associated with cardiovascular alterations.

- in a dog model of pulmonary hypertension (induced by sephadex embolization), AI-700 caused a dose-related increase in pulmonary artery pressure, including effects at doses only slightly higher than those proposed for clinical use.

- in a carotid artery injection in rats, signs of abnormal behavior occurred at AI-700 doses higher than those proposed for clinical use.

Overall, the clinical development program consisted of 11 studies that exposed 1,194 subjects to AI-700. Two phase 3 clinical studies of similar design provided the confirmatory safety and efficacy data, as outlined below.

b. Efficacy Data:

Design and study conduct:

Data from two phase 3 clinical studies were submitted to support the efficacy of AI-700. Both studies (referred to as Study 32 and Study 33) were of similar design and similar analytical methodology. Clinical sites participating in these two studies had to meet certain echocardiographic technical expectations, based upon assessments performed in a pilot study that preceded studies 32 and 33.

Studies 32 and 33 were open label (with masked central readers of images), single arm, multicenter, international studies in which "stable" patients underwent a screening echocardiogram to exclude inappropriate patients (because of certain structural or functional findings or inadequate visualization). Enrolled subjects underwent baseline ("rest") AI-700 contrast-enhanced echocardiography and SPECT (following the administration of Tc99m myocardial imaging agent). Subsequently, subjects received an injection of dipyridamole and again underwent AI-700 contrast-enhanced echocardiography and SPECT. Safety data were collected over the subsequent 72 hour period. Telephone calls were used to collect cardiac-event data at 30 (both studies) and 90 day (study 32) time points.

AI-700 was administered as an intravenous injection of 0.04 mL/kg, once during rest and once following administration of dipyridamole.

The main study outcomes related to comparisons of the AI-700 performance characteristics to those of SPECT. The study protocols noted that this comparison was appropriate because it was anticipated that AI-700 imaging could be "used as a replacement for nuclear perfusion imaging in some patients."

All AI-700 echocardiography images (wall motion with myocardial perfusion) were evaluated for myocardial defects in perfusion by three readers blinded to clinical data. SPECT images were evaluated by a single, central reader in study 32 and by a majority of three readers in study 33. The primary endpoint was considered achieved if at least two of the 3 echocardiography readers demonstrated non-inferiority to the SPECT comparator. The primary endpoint consisted of a relative risk comparison of performance characteristics between AI-700 and SPECT. The primary endpoint was evaluated in a hierarchical manner (accuracy was evaluated first, and then

if the non-inferiority endpoint was achieved, sensitivity and specificity were then evaluated similarly versus SPECT). The primary analysis was conducted on the modified intent-to-treat population (patients who received AI-700, had evaluable SPECT data and had an evaluable truth standard).

The analytical methods were modified during the conduct of the study, based upon FDA comments regarding the need to optimize the analyses to place the echocardiographic results in comparison to SPECT results, using a non-inferiority approach with pre-specified minimum performance levels for SPECT. FDA regarded this approach as maintaining consistency with the anticipated clinical use of AI-700. The choice of the non-inferority margin was to be based primarily upon clinical estimates of a reasonable margin. Specifically, for each performance characteristic, the relative ratio (echocardiography divided by SPECT) was calculated with non-inferiority being established if the lower limit of the 95% confidence interval for the ratio was ≥ 0.83. This margin of non-inferiority was less conservative than FDA had recommended.

Image assessments from echocardiography and SPECT were classified as either "disease or no disease." Similarly, truth standards resulted in patients being classified as having either "disease or no disease." The truth standard was defined by quantitative coronary arteriography plus left ventriculogram if available, or 90 day outcome as analyzed by an assessor's perspective of the "compelling" clinical data. Compelling evidence considered positive for disease was based on: confirmed myocardial infarction, subsequent cardiac death, myocardial infarction based on ECG results (Q wave), along with elevated cardiac enzymes, other evaluable rest/stress SPECT findings, or previous imaging evidence of regional wall motion abnormality on echocardiography.

The two phase 3 studies importantly differed in the types of enrolled patients (which equates to the underlying prevalence of coronary artery disease).

-in Study 32, eligible patients were those with chest pain who were also determined to be patients in clinical need of rest and stress SPECT; in this study, the prevalence of coronary artery disease was 44%, based upon the truth standard outcome.

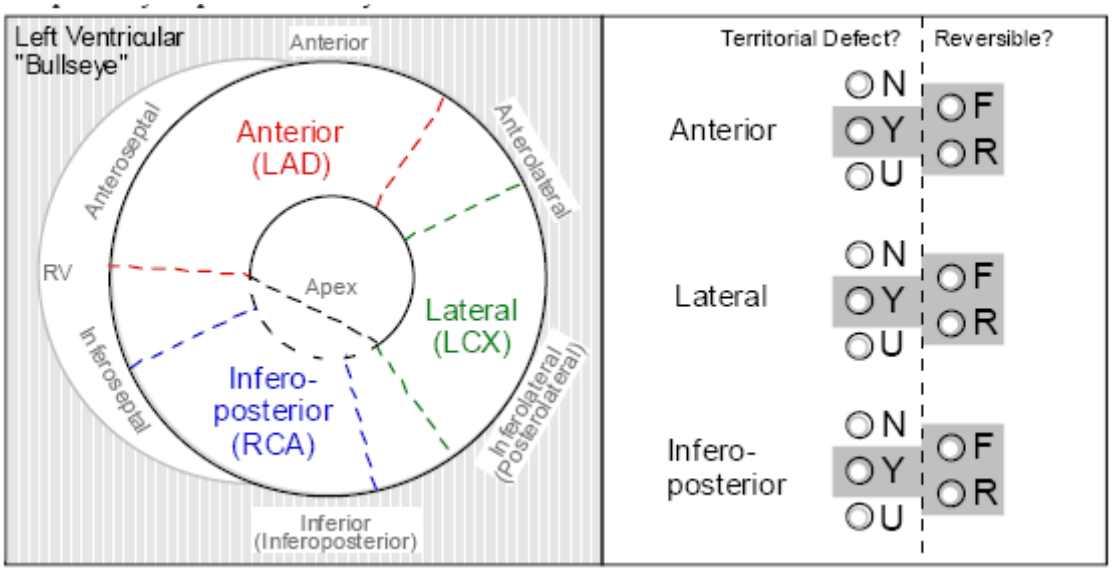
-in Study 33, eligible patients were those with chest pain who had already been determined to need coronary arteriography; the prevalence of coronary artery disease was 58%, based upon the truth standard outcome.

The AI-700 findings from Study 32 became available prior to the completion of the Study 33 echocardiographic image reads. The Study 32 results suggested that the echocardiographic readers were not interpreting the images in a manner sufficient to provide acceptable sensitivity. Hence, the on-going Study 33 echocardiographic reading process was halted and the readers were retrained (prior to any Study 33 data analyses). Results from this final read formed the Study 33 database.

Image reads:

Myocardial defect (disease) detection was based on assessment scoring maps that "collapsed" segmental scores to regional scores to territorial vascular bed scores (Figure 2). The major study outcomes were based upon overall assessments of the presence or absence of disease (not the location of perfusion defects). Supportive analyses examined defect-level outcomes.

Figure 2: Global Outcome Scoring Report



Echocardiographic readers assessed the combined myocardial wall motion/ wall thickening and contrast enhancement reads based on:

- Integrated evaluation of all contrast-enhanced views
- Change intensity of segments between rest and stress imaging
- Different intensities during different parts of cardiac cycle
- Difference enhancement intensity during different triggering intervals
- Replenishment rates after destruction pulses.

All study echocardiograms were assessed for defects (disease) by providing the following Global Outcome scores:

Y = defect	N = no defect
U = unevaluable	U = Unvisualizable
F = little or no evidence of reversible defect	R = partially or predominantly reversible defect

Angiography data were analyzed in an Independent Angiographic Core Lab by a single angiographer. Angiograms were evaluated for the diameter of stenosis, portion of myocardium subtended by stenotic vessels and regional wall abnormalities on ventriculography.

Major results:

Major baseline characteristics are summarized below.

Table 1. Major Baseline Characteristics in Study 32 and 33

Characteristic	Study 32 n = 321	Study 33 n = 457
Age (years, mean)	61	62
Male	67%	80%
Race/white	61%	79%
"Cardiovascular history"		
hyperlipidemia	73%	77%
hypertension	75%	76%
prior cardiac catheterization	60%	48%
prior CABG	17%	excluded from enrollment

CABG = coronary artery bypass grafting

Patient disposition is summarized below.

Table 2. Patient Disposition in Studies 32 and 33

group	Study 32	Study 33
Enrolled	321	457
Safety population	321	457
Modified Intent to Treat (mITT)	285	377
Excluded from mITT	36*	80
missing or ineligible SPECT	6	78
no truth standard	29	2

*one patient was recorded as enrolled twice;
mITT = received AI-700 (any part of a dose for Study 33; all of a dose for Study 33), had evaluable stress SPECT and a truth standard.

The primary endpoint component of non-inferiority for accuracy was met in both trials. The other components within the endpoint (sensitivity and specificity) were not achieved.

Since AI-700 is proposed for use as a screening tool to assist the clinician in selecting patients to undergo coronary arteriography and because AI-700 is potentially to be used as a replacement for SPECT, assessment of sensitivity is especially important. Only in Study 33 (the study where all patients were predetermined to need coronary arteriography), did AI-700 demonstrate non-inferior sensitivity to SPECT.

The major study performance characteristics (primary endpoint components) are shown below, with highlighting for individual reader outcomes that achieved non-inferiority to SPECT.

Table 3. Performance Characteristics, by Echocardiography Reader

Diagnostic Statistic	Study 32				Study 32			
	SPECT	ECHO #1	ECHO #2	ECHO #3	SPECT	ECHO #1	ECHO #2	ECHO #3
Accuracy	70 %	66 %	67 %	71 %	67 %	66 %	70 %	70 %
Ratio		0.94	0.96	1.0		0.98	1.0	1.0
Ratio CI limit		**0.86**	**0.87**	**0.93**		**0.89**	**0.96**	**0.96**
Sensitivity	78 %	77 %	57 %	50 %	61 %	73 %	68 %	73 %
Ratio		0.99	0.73	0.64		1.2	1.1	1.2
Ratio CI limit		**0.88**	0.63	0.54		**1.08**	**1.01**	**1.09**
Specificity	64 %	58 %	75 %	88 %	76 %	55 %	72 %	66 %
Ratio		0.91	1.2	1.4		0. 73	0.95	0.87
Ratio CI limit		0.78	**1.04**	**1.24**		0.62	**0.84**	0.76

Bolded portions denote successful achievement of non-inferiority (≥ 0.83 Lower limit 95% CI)
CI = confidence interval; "Echo # 1, 2, 3" refers to each individual reader

c. Safety Data:

1) Exposure:

Overall, 1,194 subjects were exposed to AI-700, as shown in table 4. These data were obtained from 11 clinical studies that enrolled healthy volunteers as well as patients.

Table 4. AI-700 Exposure

n	Subjects
1,194	Total population
911	Suspected or known CAD/proposed dose regimen
167	Co-morbidities and/or alternate dose regimens
116	Healthy volunteers

CAD = coronary artery disease

The population of 911 patients with suspected or known CAD consists of patients who received AI-700 as a component of stress testing within Study 23 (the pilot study that facilitated training of clinical sites) or Studies 32 and 33 (the phase 3 clinical studies). These patients received the proposed market dose of AI-700 (2 injections of 0.04 mL/kg) and stress testing with the use of dipyridamole. Consequently, this population of 911 patients represents the most market-applicable patient population for data analyses and is referred to as the "Phase 3 Safety Database."

As summarized below, an exploratory clinical study in healthy volunteers indicated that AI-700 administration resulted in an apparent acute inflammatory response. The extent to which this

inflammatory response contributed to the adverse reactions detected in the Phase 3 Safety Database is unclear, although the nature of several adverse events (such as hypotension, flushing, rigors) suggest a pathophysiologic analogy between the events and an acute inflammatory response.

2) Acute Inflammatory Response:

Study AI-700-04 showed that AI-700 administration initiated an acute systemic inflammatory response, based upon elevations in C3a (a biomarker of complement activation), C-reactive protein and white blood cell (WBC) count alterations.

Study AI-700-04 was designed to assess the impact of AI-700 upon acute inflammatory biomarkers as well as the immunoreactivity to the product. In Stage 1, subjects received AI-700 (no stress) and underwent follow-up laboratory evaluations. Stage 2 began one year following Stage 1 and consisted of administration of AI-700 again (rechallenge) to subjects who had a negative skin prick test to the product, along with follow-up laboratory evaluations.
Overall, 21 subjects received AI-700 in the study. Based on skin prick test and laboratory test results, no subjects evidenced an allergic or hypersensitive response in association with AI-700 rechallenge. However, changes in biomarkers of inflammation were detected in both Stage 1 and Stage 2 of the study. The following tables summarize the findings for C3a, C-reactive protein and neutrophil counts.

Table 5. C3a Concentrations (ng/mL) by Treatment Group, Stage and Timepoint*
(mean and standard deviation; normal level < 940 ng/mL)

Timepoint	AI-700		Placebo	
	Stage 1 n = 10	Stage 2 n = 7	Stage 1 n = 2	Stage 2 n = 2
- 10 minutes	31 (29)	69 (42)	3002 (4209)**	42 (14)
6 minutes	442 (451)	342 (236)	21 (4)	29 (6)
30 minutes	124 (135)	154 (58)	28 (19)	45 (11)
120 minutes	44 (22)	76 (31)	34 (21)	39 (8)

*relative to AI-700 administration;
**for unclear reasons, one placebo subject had a baseline value of 5978 recorded, all subsequent values for this subject were lower.

Table 6. C-Reactive Protein Results (mg/dL) by Treatment Group, Stage and Timepoint*
(mean and standard deviation; normal level ≤ 0.29 mg/dL)

Timepoint	AI-700		Placebo	
	Stage 1 n = 10	Stage 2 n = 7	Stage 1 n = 2	Stage 2 n = 2
- 10 minutes	0.13 (0.2)	0.14 (0.1)	0.39 (0.5)	0.20 (0.3)
120 minutes	NA	0.14 (0.1)	NA	0.18 (0.2)
24 hours	0.58 (0.3)	NA	0.24 (0.3)	NA

*relative to AI-700 administration; NA = not assessed.

During Stage 1, all 10 AI-700 subjects experienced an increase in C-reactive protein from

baseline to 24 hours post dose. Nine of the 10 subjects had increases from normal levels to higher than normal levels. No post-dose increase from baseline was seen in the placebo-treated subjects. In Stage 2, C-reactive protein levels were not assessed at 24 hours following AI-700 administration.

Routine laboratory tests were remarkable predominantly for fluctuations in WBC (specifically neutrophil) counts (Table 7). The fluctuations were characterized by a rapid (within a few minutes) decrease in counts followed by an increase in WBC counts over the subsequent hours. In general, the obtained values remained within normal ranges or minimally deviated outside the normal ranges. However, the dynamic pattern of the changes suggested that AI-700 administration was associated with inflammatory cell/hematologic bioactivity, perhaps related to complement activation.

Table 7. WBC Counts (x 10^3/mcL) by Treatment Group, Stage and Timepoint*
(mean and standard deviation; normal level ≤ 0.29 mg/dL)

Timepoint	AI-700		Placebo	
	Stage 1 n = 10	Stage 2 n = 7	Stage 1 n = 2	Stage 2 n = 2
- 10 minutes	6.0 (1.8)	6.6 (1.2)	7.3 (1.0)	5.7 (2.3)
6 minutes	3.9 (1.6)	5.5 (1.1)	6.6 (1.0)	4.8 (1.1)
30 minutes	5.6 (1.5)	5.8 (1.3)	6.4 (1.8)	6.0 (2.3)
1 hour	5.4 (1.9)	5.4 (1.2)	6.7 (1.4)	5.8 (1.8)
2 hours	5.6 (2.4)	5.5 (1.5)	6.9 (0.1)	7.5 (3.9)
8 hours	9.3 (3.5)	NA	8.4 (0.2)	NA
24 hours	5.9 (1.8)	NA	6.8 (0.3)	NA

*relative to AI-700 administration; NA = not assessed

Other laboratory values obtained in the study, including serum tryptase levels, total C3 and clinical chemistry, showed no remarkable alterations. In general, the biomarker changes (C3a, C-reactive protein, WBC counts) did not manifest as clinically symptomatic reactions within the study subjects.

3) Adverse Events:

Overview:

Summarized below are the major safety findings. No deaths were reported following AI-700 administration. However, numerous adverse events were reported.

Because the all serious adverse events and the vast majority of adverse events occurred in the "Phase 3 Safety Database" (Study 32 and 33, as well as Study 23 that "qualified" sites) of 911 subjects, this database provides the most cogent data source.

The distribution of major adverse events within the Phase 3 Safety Database is shown below. Overall, 1879 adverse events were experienced by 652 patients. Of the 1870 adverse events, severity was assessed as:

- mild in 77%
- moderate in 20%
- severe in ~ 2%

Adverse events were graded for severity using the following criteria:

- mild: did not interfere with daily activities
- moderate: interfered with daily activities
- severe: prevented daily activities

Table 8. Adverse Events in Phase 3 Safety Database

Event	Patients n = 911
Any adverse event	652 (72%)
Any adverse event requiring treatment	172 (19%)
Any serious adverse event	11 (1%)
Any adverse event causing AI-700 discontinuation	16 (2%)
Any adverse event causing AI-700 dose interruption or dose adjustment	7 (1%)

The interpretation of adverse events in the Phase 3 Safety Database involves consideration of the confounding impact of administration of dipyridamole and the response to this stress agent.

Serious adverse events:

Within the Phase 3 Safety Database, 11 patients experienced serious adverse events, as summarized in the table, below. To facilitate interpretation of the timing of events, the study procedures are shown in Figure 3.

Figure 3. Study Procedures:

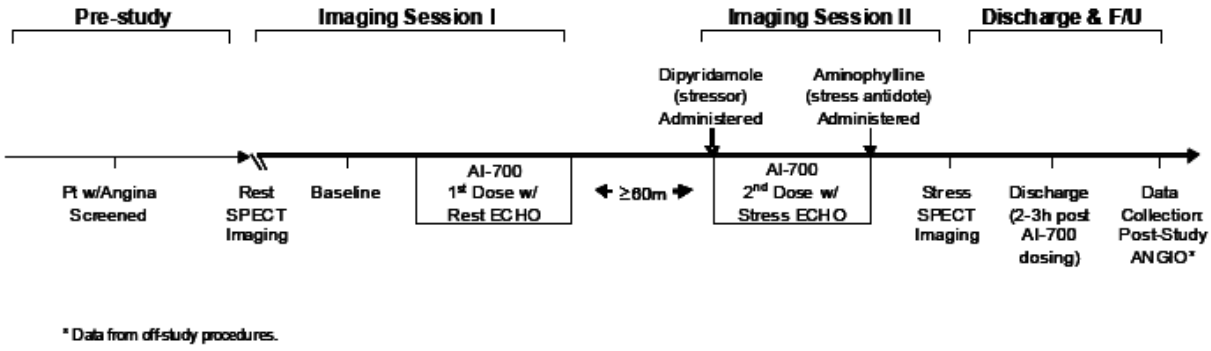


Table 9. Serious Adverse Events (SAEs)

Patient Age & Gender	Events	Timing	Resolution & Treatment
76/Female	vertigo & hypertension	within 30 min after first AI-700 dose	vertigo resolved without treatment; hypertension resolved with metoprolol
79/Male	Syncope, with junctional rhythm, pulse of 34 and SBP of 60 mm Hg	within 30 min after first AI-700 dose	bradycardia and hypotension treated with atropine and polygeline
54/Female	Mental status change with lethargy and unresponsiveness	within 30 min after second AI-700 dose	reaction was treated with steroids, benadryl and saline
47/Male	Hospitalized for eye pain, blurred vision	~ 18 hours following second AI-700 dose	resolved without treatment
63/Male	Hospitalized for fever, chills, dizziness, headache, flushing	~ 3 hours following second AI-700 dose	resolved with paracetamol
69/Male	Hospitalized with bronchospasm	~ 1 hour after second AI-700 dose	resolved with aminophylline, steroid therapy and chlorpheniramine
72/Female	Syncope, junctional rhythm, pulse of 39, undetectable BP	within 30 min after first AI-700 dose	resolved with atropine, saline
59/Female	Syncope, flushing with decrease in pulse from 80 to 55, no hypotension	within 30 min after first AI-700 dose	resolved with atropine
76/Male	Myocardial infarction	~ 2 days after second AI-700 dose	resolved with therapy
80/Female	Myocardial infarction	~ 2 days after second AI-700 dose	resolved with therapy
64/Male	Chest pain	within 30 min after second AI-700 dose	resolved with nitroglycerin, tirofiban

Within the entire AI-700 exposure database, 17 subjects experienced adverse events that were notable enough to cause permanent AI-700 discontinuation (1 healthy volunteer subject and 16 patients within the Phase 3 Safety Database). These subjects are summarized below. Four of these patients had the events classified as SAE.

Table 10. Adverse Events Resulting in Permanent Discontinuation of AI-700

Patient Age & Gender	Event(s)	Timing	Resolution & Treatment
36/Female	chest pain & tachycardia	during second AI-700 dose	resolved without treatment
76/Female*	vertigo, hypertension, headache	after first AI-700 dose	only hypertension required treatment
79/Male*	syncope	after first AI-700 dose	resolved with treatment
68/Male	hypotension	after first AI-700 dose	resolved without treatment
56/Female	headache & rigors	after first AI-700 dose	resolved without treatment
64/Female	flushing, lacrimation, wheezing, coughing	after first AI-700 dose	only treatment needed for cough
68/Male	sweating, hypotension, tachypnea, vertigo	after first AI-700 dose	resolved with promethazine treatment
66/Female	dizziness & hypotension	during the second AI-700 dose	only hypotension required treatment
69/Male	rigors	after first AI-700 dose	resolved with treatment
65/Female	hypotension	during second AI-700 dose	resolved with treatment
62/Female	hypersensitivity manifest as hypotension	after first AI-700 dose	resolved with treatment
72/Female*	syncope,	after first AI-700 dose	resolved with treatment
61/Female	allergic reaction manifest as decrease in oxygen saturation from 99% to 89%	after first AI-700 dose	resolved with treatment
59/Female*	syncope	during first AI-700 dose	resolved with treatment
73/Female	angina with ST-segment depression	after first AI-700 dose	resolved with treatment
66/Male	cyanosis & decreased oxygen saturation	after first AI-700 dose	resolved with treatment
73/Female	weakness, dizziness, diarrhea, hypotension, bradycardia, mild QT	after first AI-700 dose	resolved with treatment

	prolongation		

*this patient was also classified as having a serious adverse event

Overall, seven patients experienced adverse events that prompted either temporary interruption of the AI-700 dose injection or adjustment of the dosing regimen. These are summarized below.

Table 11. Adverse Events Resulting in Temporary AI-700 Interruption or Dose Adjustment

Patient Age & Gender	Event(s)	Timing	Resolution & Treatment
64/Male	chest pain	temporary interruption of second AI-700 dose	resolved with aminophylline
50/Male	"cold sweat"	temporary interruption of second AI-700 dose	resolved without treatment
78/Male	hypotension	adjustment of AI-700 dose during second administration	resolved with intravenous saline
64/Female	bradycardia, nausea, vomiting	temporary interruption of second AI-700 dose	resolved with aminophylline
64/Male	nausea, headache, decreased blood pressure, dizziness	temporary interruption of second AI-700 dose	resolved without treatment
58/Male	chest pain, headache, ear, neck and throat pain	temporary interruption of second AI-700 dose	resolved without treatment
52/Male	chest pain	temporary interruption of second AI-700 dose	resolved with aminophylline, nitroglycerin

The general pattern of important adverse events is exemplified by the following narratives:

1) A 62 year old female with a history of myocardial infarction, diabetes mellitus, hypertension, atypical angina and hyperlipidemia received her first AI-700 dose and approximately five minutes later her blood pressure had decreased to 85/60. The patient was treated with hydrocortisone, acetaminophen and chlorpheniramine and her blood pressure returned to near baseline (128/89 by 60 minutes post-dose). AI-700 was permanently discontinued and the reactions was classified as "severe hypersensitivity" (non-serious).

2) A 72 year old female with a history of angina, myocardial infarction, hyperlipidemia, hypothyroidism, osteoporosis and recurrent urinary tract infections experienced a sensation of flushing and dizziness less than one minute following initiation of her first AI-700 dose. Over the subsequent 20 minutes she developed a junctional cardiac rhythm with a ventricular rate of 39 and signs of "near syncope." Her blood pressure was not measurable by auscultation.

Intravenous atropine and saline were administered and cardiac rhythm returned to sinus with a heart rate of 76 and resolution of syncope symptoms. She was hospitalized for observation and the event was recorded as a serious adverse event that resulted in permanent AI-700 discontinuation.

3) A 61 year old female with a history of diabetes, osteoarthritis, hypertension, exertional angina, hyperlipidemia, recurrent laryngitis, retinopathy and hypothyroidism experienced a decrease in cutaneous oxygen saturation from 99% to 89% within three minutes following AI-700 injection. She also experienced a slight increase in temperature (from 37.2 to 37.9ºC). She was treated with oxygen and improved to no longer need oxygen. The event was classified as an allergic reaction (non-serious) and prompted permanent AI-700 discontinuation.

4) A 66 year old male (former smoker) with a history of angina, congestive heart failure and lumbar fracture experienced cutaneous oxygen desaturation within 10 minutes after an initial AI-700 dose. Oxygen saturation decreased from 95% at baseline to 83% and was associated with flushing and acrocyanosis. Supplemental oxygen was administered and the oxygen desaturation resolved over the subsequent hour. The event was classified as non-serious but prompted permanent AI-700 discontinuation.

5) A 73 year old female with a history of diabetes, hypertension, angina, stroke, hyperlipidemia, hypothyroidism and Meniere's disease developed nausea, dizziness and other malaise symptoms within 10 minutes following the start of an initial AI-700 dose. Shortly following the onset of the symptoms, the investigator detected bradycardia, hypotension and heart block that quickly resolved with no specific treatment. None of the events required treatment. The events were recorded as non-serious but prompted permanent AI-700 discontinuation.

A summary of the most common adverse events is duplicated in the sponsor's table summary, as follows on the next page.

Table 12. Number of Patients with Adverse Events Reported in ≥ 1% of Patients, by Onset Time (Phase 3 Safety Database)

	Onset Time				
Preferred Term	1st Dose to DIPY N = 911 n (%)	DIPY to 2nd Dose N = 894 n (%)	2nd Dose to Discharge N = 894 n (%)	Discharge to Follow-up N = 911 n (%)	Total N = 911 n (%)
Number of Patients with ≥1 AE	105 (11.5)	391 (43.7)	402 (45.0)	151 (16.6)	652 (71.6)
Total Number of AEs	179	672	768	260	1879
Headache	27 (3.0)	167 (18.7)	128 (14.3)	28 (3.1)	317 (34.8)
Chest pain	2 (0.2)	60 (6.7)	41 (4.6)	10 (1.1)	101 (11.1)
Nausea	5 (0.5)	34 (3.8)	45 (5.0)	13 (1.4)	95 (10.4)
Flushing	17 (1.9)	52 (5.8)	18 (2.0)	3 (0.3)	85 (9.3)
Chest discomfort	4 (0.4)	45 (5.0)	22 (2.5)	4 (0.4)	68 (7.5)
Leukocytosis*	1 (0.1)	0	57 (6.4)	1 (0.1)	59 (6.5)
Dizziness	7 (0.8)	27 (3.0)	16 (1.8)	5 (0.5)	52 (5.7)
Neutrophil count increased*	2 (0.2)	0	46 (5.1)	3 (0.3)	51 (5.6)
Feeling hot	6 (0.7)	29 (3.2)	6 (0.7)	4 (0.4)	45 (4.9)
Dyspnea	4 (0.4)	20 (2.2)	16 (1.8)	3 (0.3)	40 (4.4)
Rigors	10 (1.1)	2 (0.2)	16 (1.8)	13 (1.4)	39 (4.3)
ECG ST segment depression	1 (0.1)	18 (2.0)	15 (1.7)	0	33 (3.6)
Hypotension	9 (1.0)	4 (0.4)	18 (2.0)	1 (0.1)	31 (3.4)
Abdominal pain	0	6 (0.7)	11 (1.2)	5 (0.5)	22 (2.4)
Lymphocyte count decreased*	2 (0.2)	0	20 (2.2)	0	22 (2.4)
Pain in extremity	1 (0.1)	9 (1.0)	9 (1.0)	1 (0.1)	20 (2.2)
Neck pain	0	12 (1.3)	5 (0.6)	1 (0.1)	18 (2.0)
Pyrexia	1 (0.1)	0	6 (0.7)	12 (1.3)	19 (2.1)
Fatigue	1 (0.1)	2 (0.2)	4 (0.4)	8 (0.9)	15 (1.6)
Hyperhidrosis	5 (0.5)	3 (0.3)	4 (0.4)	5 (0.5)	17 (1.9)
Abdominal pain upper	1 (0.1)	9 (1.0)	5 (0.6)	0	15 (1.6)
Throat tightness	0	7 (0.8)	3 (0.3)	1 (0.1)	10 (1.1)
Dysgeusia	3 (0.3)	6 (0.7)	3 (0.3)	1 (0.1)	13 (1.4)
Monocyte count decreased*	1 (0.1)	0	12 (1.3)	0	13 (1.4)
Vomiting	1 (0.1)	3 (0.3)	4 (0.4)	4 (0.4)	12 (1.3)
Hot flush	0	6 (0.7)	5 (0.6)	1 (0.1)	12 (1.3)
Pain in jaw	1 (0.1)	6 (0.7)	6 (0.7)	0	12 (1.3)
Tremor	1 (0.1)	0	9 (1.0)	1 (0.1)	11 (1.2)
Paresthesia	1 (0.1)	6 (0.7)	3 (0.3)	0	10 (1.1)
WBC count increased*	0	0	11 (1.2)	0	11 (1.2)
Burning sensation	0	7 (0.8)	3 (0.3)	0	10 (1.1)
Injection site burning	1 (0.1)	7 (0.8)	2 (0.2)	0	10 (1.1)
Angina pectoris	1 (0.1)	4 (0.4)	2 (0.2)	2 (0.2)	9 (1.0)
Myalgia	1 (0.1)	1 (0.1)	0	7 (0.8)	9 (1.0)

* Post-dose clinical laboratory samples were collected only at Discharge and Follow-up.

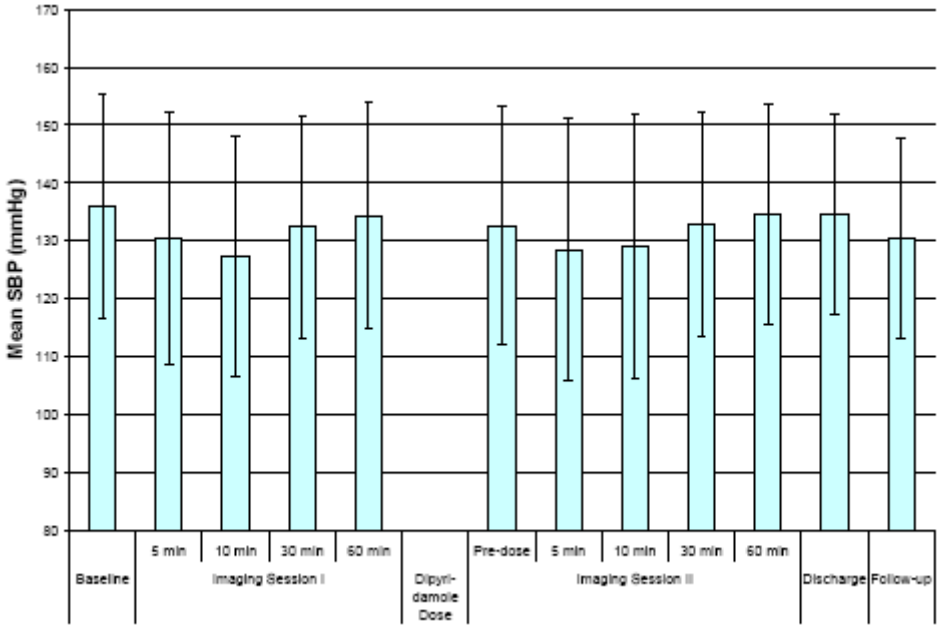
Clinical Laboratory and Vital Sign Changes

Clinical laboratory findings were most notable for alterations in WBC counts (increases in neutrophils and decreases in lymphocytes) following AI-700 administration. The shifts were generally experienced by approximately 20% of the patients and were evident at 2 to 3 hours following AI-700 injection. The hematologic changes largely resolved by 72 hours following AI-700 injection. Most alterations were assessed as "mild to moderate" grade alterations.

Vital sign alterations for the groups of patients receiving AI-700 were generally not clinically

significant, although the pattern of blood pressure alteration was suggestive of a slight decrease following AI-700 administration as shown in Figure 4.

Figure 4. Mean Systolic Blood Pressure (± SC) over Time (Studies 32 and 33)



4. FDA Guidance Document Bibliography

1. Guidance for Industry: Developing Medical Imaging Drug and Biological Products/Part 1: Conducting Safety Assessments

2. Guidance for Industry: Developing Medical Imaging Drug and Biological Products/Part 2: Clinical Indications

3. Guidance for Industry: Developing Medical Imaging Drug and Biological Products/Part 3: Design, Analysis and Interpretation of Clinical Studies

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 1: Conducting Safety Assessments

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 1: Conducting Safety Assessments

Additional copies of this Guidance are available from:

Division of Drug Information HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Rockville, MD 20857
(Phone 301-827-4573)
Internet: http://www.fda.gov/cder/guidance/index.htm.

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
Internet: http://www.fda.gov/cber/guidelines.htm.
Mail: the Voice Information System at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Table of Contents

I. INTRODUCTION **1**

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS **2**

A. Contrast Agents **2**

B. Diagnostic Radiopharmaceuticals **2**

III. GENERAL CONSIDERATIONS FOR SAFETY ASSESSMENTS OF MEDICAL IMAGING AGENTS **3**

A. Medical Imaging Agent Characteristics Relevant to Safety **3**

1. Mass Dose *4*
2. Route of Administration *4*
3. Frequency of Use *4*
4. Biological, Physical, and Effective Half-Lives *5*

B. Performance of Nonclinical Safety Assessments **5**

1. Nonclinical Safety Assessments for Nonbiological Drug Products *5*
a. Timing of Nonclinical Studies Submitted to an IND Application 5
b. Contrast Agents 6
c. Diagnostic Radiopharmaceuticals (Nonbiological Products) 8
2. Nonclinical Safety Assessments for Biological Products *8*

IV. CLINICAL SAFETY ASSESSMENTS **9**

A. Group 1 and 2 Medical Imaging Agents **9**

1. Group 1 Medical Imaging Agents *9*
2. Group 2 Medical Imaging Agents *10*

B. Considerations For Groups 1 or 2 ..... **11**

1. Safety-Margin Considerations *11*
a. Results of nonclinical studies 11
b. Results of initial human experience 13
2. Clinical Use Considerations *14*

C. Radiation Safety Assessment for All Diagnostic Radiopharmaceuticals **14**

1. General Considerations *14*
2. Calculation of Radiation Absorbed Dose to the Target Organs or Tissues *15*
3. Maximum Radiation Absorbed Dose *16*

Guidance for Industry[1]
Developing Medical Imaging Drug and Biological Products
Part 1: Conducting Safety Assessments

This guidance represents 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

This guidance is one of three guidances intended to assist developers of medical imaging drug and biological products (*medical imaging agents*) in planning and coordinating their clinical investigations and preparing and submitting investigational new drug applications (INDs), new drug applications (NDAs), biologics license applications (BLAs), abbreviated NDAs (ANDAs), and supplements to NDAs or BLAs. The three guidances are*: Part 1: Conducting Safety Assessments*; *Part 2: Clinical Indications*; and *Part 3: Design, Analysis, and Interpretation of Clinical Studies.*

Medical imaging agents generally are governed by the same regulations as other drug and biological products. However, because medical imaging agents are used solely to diagnose and monitor diseases or conditions as opposed to treat them, development programs for medical imaging agents can be tailored to reflect these particular uses. Specifically, this guidance discusses our recommendations on conducting safety assessments of medical imaging agents.

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 guidances means that something is suggested or recommended, but not required.

[1] This guidance has been prepared by the Division of Medical Imaging and Radiopharmaceutical Drug Products and the Office of Therapeutics Research and Review in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration.

A glossary of common terms used in diagnostic medical imaging is provided at the end of this document.

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS

This guidance discusses medical imaging agents that are administered in vivo and are used for diagnosis or monitoring with a variety of different modalities, such as radiography, computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), and radionuclide imaging. The guidance is not intended to apply to the development of in vitro diagnostic or therapeutic uses of these agents.[2]

Medical imaging agents can be classified into at least two general categories, contrast agents and diagnostic radiopharmaceuticals.

A. Contrast Agents

As used in this guidance, a *contrast agent* is a medical imaging agent used to improve the visualization of tissues, organs, and physiologic processes by increasing the relative difference of imaging signal intensities in adjacent regions of the body. Types of contrast agents include, but are not limited to, (1) iodinated compounds used in radiography and CT; (2) paramagnetic metallic ions (such as ions of gadolinium, iron, and manganese) linked to a variety of molecules and microparticles (such as superparamagnetic iron oxide) used in MRI; and (3) microbubbles, microaerosomes, and related microparticles used in diagnostic ultrasonography.

B. Diagnostic Radiopharmaceuticals

As used in this guidance, a *diagnostic radiopharmaceutical* is (1) an article that is intended for use in the diagnosis or monitoring of a disease or a manifestation of a disease in humans and that exhibits spontaneous disintegration of unstable nuclei with the emission of nuclear particles or photons or (2) any nonradioactive reagent kit or nuclide generator that is intended to be used in

---

[2] The guidance is not intended to apply to the development of research drugs that do not provide direct patient benefit with respect to diagnosis, therapy, prevention, or prognosis, or other clinically useful information. These include radioactive drugs for research that are used in accordance with 21 CFR 361.1. Section 361.1(a) states that radioactive drugs (defined in 21 CFR 310.3(n)) are generally recognized as safe and effective when administered under specified conditions to human research subjects in the course of a project intended to obtain basic information about the metabolism of a radioactively labeled drug or about human physiology, pathophysiology, or biochemistry. However, if a radioactive drug is used for immediate therapeutic, diagnostic, or similar purposes or to determine the safety and effectiveness of the drug in humans, or if the radioactive drug has a pharmacological effect in the human body, an IND is required. FDA is developing a guidance on determining when research with radioactive drugs may be conducted under § 361.1.

The Agency recognizes the potential of imaging agents as research tools for aiding the development of therapeutic drugs, and some of the principles of the guidance may be applicable to such research. Sponsors of such imaging research agents are urged to contact the Division of Medical Imaging and Radiopharmaceutical Drug Products for advice on development of the imaging research agent.

the preparation of such an article.[3] As stated in the preamble to FDA's proposed rule on Regulations for In Vivo Radiopharmaceuticals Used for Diagnosis and Monitoring, the Agency interprets this definition to include articles that exhibit spontaneous disintegration leading to the reconstruction of unstable nuclei and the subsequent emission of nuclear particles or photons (63 FR 28301 at 28303; May 22, 1998).

Diagnostic radiopharmaceuticals are generally radioactive drug or biological products that contain a radionuclide that typically is linked to a ligand or carrier.[4] These products are used in nuclear medicine procedures, including planar imaging, single photon emission computed tomography (SPECT), positron emission tomography (PET), or in combination with other radiation detection probes.

Diagnostic radiopharmaceuticals used for imaging typically have two distinct components.

- A radionuclide that can be detected in vivo (e.g., technetium-99m, iodine-123, indium-111).

  The radionuclide typically is a radioactive atom with a relatively short physical half-life that emits radioactive decay photons having sufficient energy to penetrate the tissue mass of the patient. These photons can then be detected with imaging devices or other detectors

- A nonradioactive component to which the radionuclide is bound that delivers the radionuclide to specific areas within the body.

  This nonradionuclidic portion of the diagnostic radiopharmaceutical often is an organic molecule such as a carbohydrate, lipid, nucleic acid, peptide, small protein, or antibody.

As technology advances, new products may emerge that do not fit into these traditional categories (e.g., agents for optical imaging, magnetic resonance spectroscopy, combined contrast and functional imaging). It is anticipated, however, that the general principles discussed here could apply to these new diagnostic products. Developers of these products should contact the appropriate reviewing division for advice on product development.

III. GENERAL CONSIDERATIONS FOR SAFETY ASSESSMENTS OF MEDICAL IMAGING AGENTS

A. Medical Imaging Agent Characteristics Relevant to Safety

[3] 21 CFR 315.2 and 601.31.

[4] In this guidance, the terms *ligand* and *carrier* refer to the entire nonradionuclidic portion of the diagnostic radiopharmaceutical.

112 The following sections discuss the special characteristics of a medical imaging agent that can
113 lead to a more focused safety evaluation. Characteristics include its radiation absorbed dose,
114 mass dose, route of administration, frequency of use, biodistribution, and biological, physical,
115 and effective half-lives in the serum, the whole body, and critical organs.[5]
116
117 *1. Mass Dose*
118
119 Some medical imaging agents can be administered at low mass doses. For example, the mass
120 dose of a single administration of a diagnostic radiopharmaceutical can be small because device
121 technologies can typically detect relatively small amounts of a radionuclide (e.g.,
122 radiopharmaceuticals for myocardial perfusion imaging). When a medical imaging agent is
123 administered at a mass dose that is at the low end of the dose-response curve, dose-related
124 adverse events are less likely to occur.
125
126 *2. Route of Administration*
127
128 Some medical imaging agents are administered by routes that decrease the likelihood of systemic
129 adverse events. For example, medical imaging agents that are administered as contrast media for
130 radiographic examination of the gastrointestinal tract (e.g., barium sulfate) can be administered
131 orally, through an oral tube, or rectally. In patients with normal gastrointestinal tracts, many of
132 these products are not absorbed, so systemic adverse events are less likely to occur. In general,
133 nonradiolabeled contrast agents pose safety issues similar to therapeutic drugs because of the
134 inherently large amounts needed for administration. Therefore, nonradiolabeled drugs generally
135 should be treated like therapeutic agents for the purpose of conducting clinical safety
136 assessments.
137
138 *3. Frequency of Use*
139
140 Many medical imaging agents, including both contrast agents and diagnostic
141 radiopharmaceuticals, are administered infrequently or as single doses. Accordingly, adverse
142 events that are related to long-term use or to accumulation are less likely to occur with these
143 agents than with agents that are administered repeatedly to the same patient. Therefore, the
144 nonclinical development programs for such single-use products usually can omit long-term (i.e.,
145 3 months' duration or longer), repeat-dose safety studies. In clinical settings where it is possible
146 that the medical imaging agent will be administered to a single patient repeatedly (e.g., to
147 monitor disease progression), we recommend that repeat-dose studies (of 14 to 28 days'
148 duration) be performed to assess safety.
149
150 Biological medical imaging agents are frequently immunogenic, and the development of
151 antibodies after intermittent, repeated administration can alter the pharmacokinetics,
152 biodistribution, safety, and/or imaging properties of such agents and, potentially, of
153 immunologically related agents. We recommend that studies in which repeat dosing of a
154 biological imaging agent is planned incorporate pharmacokinetic data, human anti-mouse

[5] See also 21 CFR 315.6 on evaluation of safety. When a medical imaging agent does not possess any of these special characteristics, as described in section III.A.1-4, complete standard safety assessments should be performed.

155 antibody (HAMA), human anti-humanized antibody (HAHA), or human anti-chimeric antibody
156 (HACA) levels as well as whole body biodistribution imaging to assess for alterations in the
157 biodistribution of the imaging agent following repeat dosing. Studies of immunogenicity in
158 animal models are generally of limited value. Therefore, we recommend that human clinical
159 data assessing the repeat use of a biological imaging agent be obtained prior to application for
160 licensure of such an agent.
161
162 *4. Biological, Physical, and Effective Half-Lives*
163
164 Diagnostic radiopharmaceuticals often use radionuclides with short physical half-lives or that are
165 excreted rapidly. The biological, physical, and effective half-lives of diagnostic
166 radiopharmaceuticals are incorporated into radiation dosimetry evaluations[6] that require an
167 understanding of the kinetics of the distribution and excretion of the radionuclide and its mode of
168 decay. We recommend that biological, physical, and effective half-lives be considered in
169 planning appropriate safety and dosimetry evaluations of diagnostic radiopharmaceuticals.
170
171 **B. Performance of Nonclinical Safety Assessments**
172
173 We recommend that the nonclinical development strategy for an agent be based on sound
174 scientific principles, the agent's unique chemistry (including, for example, those of its
175 components, metabolites, and impurities), and the agent's intended use. Because each product is
176 unique, we encourage sponsors to consult with us before submitting an IND application and
177 during product development. The number and types of nonclinical studies recommended would
178 depend in part on the phase of development, what is known about the agent or its pharmacologic
179 class, its proposed use, and the indicated patient population. If you determine that nonclinical
180 pharmacology or toxicology studies are not needed, we are prepared to grant a waiver under
181 21 CFR 312.10 if you provide adequate justification.
182
183 In the discussion that follows, a distinction is made between drug products and biological
184 products. Existing specific guidance for biological products is referenced but not repeated here
185 (see section III.B.2).
186
187 *1. Nonclinical Safety Assessments for Nonbiological Drug Products*
188
189 a. Timing of Nonclinical Studies Submitted to an IND Application
190
191 We recommend that nonclinical studies be timed so that they help facilitate the
192 timely conduct of clinical trials (including appropriate safety monitoring based on
193 findings in nonclinical studies) and to reduce the unnecessary use of animals and

[6] *Biological half-life* is the time needed for a human or animal to remove, by biological elimination, half of the amount of a substance that has been administered. *Effective half-life* is the time needed for a radionuclide in a human or animal to decrease its activity by half as a combined result of biological elimination and radioactive decay. *Physical half-life* is the time needed for half of the population of atoms of a particular radioactive substance to disintegrate to another nuclear form.

194 other resources.[7] The recommended timing of nonclinical studies for medical
195 imaging drugs is summarized in Table 1.
196
197 b. Contrast Agents
198
199 Because of the characteristics of contrast drug products (e.g., variable biologic
200 half-life) and the way they are used, we recommend that nonclinical safety
201 evaluations of such drug products be made more efficient with the following
202 modifications:
203
204 • Long-term (i.e., greater than 3 months), repeat-dose toxicity studies in animals
205 usually can be omitted. (Exceptions are products with long residence time,
206 e.g., > 90 days.)
207
208 • Long-term rodent carcinogenicity studies usually can be omitted.[8]
209
210 • Reproductive toxicology studies required under § 312.23(a)(8)(ii)(*a*) often can
211 be limited to an evaluation of embryonic and fetal toxicities in rats and rabbits
212 and to evaluations of reproductive organs in other short-term toxicity studies.[9]
213 If you determine that such reproductive studies are not needed, we are
214 prepared to grant a waiver under § 312.10 if you provide adequate
215 justification.
216
217 We recommend that studies be conducted to address the effects of large mass
218 dose and volume (especially for iodinated contrast materials administered
219 intravenously); osmolality effects; potential transmetalation of complexes of
220 gadolinium, manganese, or iron (generally MRI drugs); potential effects of tissue
221 or cellular accumulation on organ function (particularly if the drug is intended to
222 image a diseased organ system); and the chemical, physiological, and physical
223 effects of ultrasound microbubble drugs (e.g., coalescence, aggregation,
224 margination, and cavitation).

---

[7] See the guidance *M3 Nonclinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals*. This and all other guidances cited in this document are available at FDA's Web site at http://www.fda.gov/cder/guidance/index.htm.

[8] Circumstances in which carcinogenicity testing may be recommended are summarized in the guidance *S1A The Need for Long-Term Rodent Carcinogenicity Studies of Pharmaceuticals*.

[9] See the guidance *S5A Detection of Toxicity to Reproduction for Medicinal Products* and *S5B Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility*.

225 **Table 1: Timing of Nonclinical Studies for Nonbiological Products Submitted to an IND**
226

Study Type	Before Phase 1	Before Phase 2	Before Phase 3	Before NDA
Safety pharmacology	Major organs,[(a)] and organ systems the drug is intended to visualize			
Toxicokinetic pharmacokinetic	See ICH guidances			
Expanded single-dose toxicity	Expanded acute single dose [(b)]			
Short-term (2 to 4 weeks) multiple dose toxicity		Repeat-dose toxicity		
Special toxicology	Conduct as necessary based on route-irritancy, blood compatibility, protein flocculation, misadministration, extravasation			
Radiation dosimetry	If applicable			
Genotoxicity	In vitro [(d)]	Complete standard battery		
Immunotoxicity			May be needed based on molecular structure, biodistribution pattern, class concern, or clinical or nonclinical signal	
Reproductive and developmental toxicity			Needed or waiver obtained [(d)]	
Drug interaction				As needed
Other based on data results				As needed

227 (a) See the guidances *S7A Safety Pharmacology Studies for Human Pharmaceutical* and *S7B Safety Pharmacology*
228 *Studies for Assessing the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human*
229 *Pharmaceuticals* (note that *S7B* allows for phase evaluation of the required studies).
230 (b) See the guidance *Single Dose Acute Testing for Pharmaceuticals*.
231 (c) When repeat-dose toxicity studies have been performed, but single-dose toxicology studies have not, dose
232 selection for initial human studies will likely be based on the results of the no-adverse-effect level (NOAEL)
233 obtained in the repeat-dose study. The likely result will be a mass dose selection for initial human administration
234 that is lower than if the dose selection had been based on the results of acute, single-dose toxicity studies.
235 (d) See radiopharmaceutical discussion in section III.B.1.c of this document.
236

237 c. Diagnostic Radiopharmaceuticals (Nonbiological Products)
238
239 Because of the characteristics of diagnostic radiopharmaceuticals and the way
240 they are used, we recommend that nonclinical safety evaluations of these drugs be
241 made more efficient by the following modifications:
242
243 • Long-term, repeat-dose toxicity studies in animals typically can be omitted.
244
245 • Long-term rodent carcinogenicity studies typically can be omitted.
246
247 • Reproductive toxicology studies can be waived when adequate scientific
248 justification is provided.[10]
249
250 • Genotoxicity studies should be conducted on the nonradioactive component
251 because the genotoxicity of the nonradioactive component should be
252 identified separately from that of the radionuclide. Genotoxicity studies can
253 be waived if adequate scientific justification is provided.[11]
254
255 We recommend that special safety considerations for diagnostic
256 radiopharmaceuticals include verification of the mass dose of the radiolabeled
257 and unlabeled moiety; assessment of the mass, toxic potency, and receptor
258 interactions for any unlabeled moiety; assessment of potential pharmacologic
259 or physiologic effects due to molecules that bind with receptors or enzymes;
260 and evaluation of all components in the final formulation for toxicity (e.g.,
261 excipients, reducing drugs, stabilizers, anti-oxidants, chelators, impurities, and
262 residual solvents). We recommend that the special safety considerations
263 include an analysis of particle size (for products containing particles) and an
264 assessment of instability manifested by aggregation or precipitation. We also
265 recommend that an individual component be tested if specific toxicological
266 concerns are identified or if toxicological data for that component are lacking.
267 However, if toxicological studies are performed on the combined components
268 of a radiopharmaceutical and no significant toxicity is found, toxicological
269 studies of individual components are seldom required.
270
271 *2. Nonclinical Safety Assessments for Biological Products*
272
273 Many biological products raise relatively distinct nonclinical issues such as immunogenicity and
274 species specificity. We recommend the following Agency documents be reviewed for guidance
275 on the preclinical evaluation of biological medical imaging agents:
276
277 • *S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals*

[10] See footnote 11.

[11] See guidances *S2A Specific Aspects of Regulatory Genotoxicity Tests for Pharmaceuticals* and *S2B Genotoxicity: A Standard Battery for Genotoxicity Testing of Pharmaceuticals*.

278
279 • *Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products*
280 *for Human Use*
281
282 Sponsors are encouraged to consult with the appropriate reviewing division for additional
283 information when needed.
284
285
286 **IV. CLINICAL SAFETY ASSESSMENTS**
287
288 Under section 505(d) of the Federal Food, Drug, and Cosmetic Act (the Act) (21 U.S.C. 355(d)),
289 FDA cannot approve a new drug application (NDA) unless it contains adequate tests
290 demonstrating whether the proposed drug product is safe for use under the conditions prescribed,
291 recommended, or suggested in its proposed labeling.[12] All drugs have risks, including risks
292 related to the intrinsic properties of the drug, the administration process, the reactions of the
293 patient, and incorrect diagnostic information. Incorrect diagnostic information includes
294 inaccurate structural, functional, physiological, or biochemical information; false positive or
295 false negative diagnostic determinations; and information leading to inappropriate decisions in
296 diagnostic or therapeutic management. Even if risks are found to be small, all drug development
297 programs must also obtain evidence of drug effectiveness under section 505 of the Act.
298 Although it has been suggested that a demonstration of effectiveness not be required for *safer*
299 drugs, this statutory requirement cannot be waived. FDA weighs the benefits and risks of each
300 proposed drug product when making its decision about whether to approve a marketing
301 application (e.g., an NDA or BLA).
302
303 **A. Group 1 and 2 Medical Imaging Agents**
304
305 The special characteristics of medical imaging agents may allow for a more efficient clinical
306 safety program. This guidance describes two general categories for medical imaging agents:
307 Group 1 and Group 2. The extent of clinical safety monitoring and evaluation that we
308 recommend differs for these two categories. Generally, a less extensive clinical safety
309 evaluation is appropriate for Group 1 agents. Conversely, we recommend that Group 2 agents
310 undergo standard clinical safety evaluations in clinical trials throughout their development.
311 These different groups have been conceived to help drug sponsors identify and differentiate
312 those characteristics that are of greatest interest to the Agency in assessing the potential safety of
313 a medical imaging agent.
314
315 FDA anticipates that it can assess which agents are Group 1 agents based on the safety-margin
316 criteria from animal studies and initial human trials completed at the end of Phase 1.
317
318 *1. Group 1 Medical Imaging Agents*
319
320 For purposes of this guidance, a Group 1 medical imaging agent generally exhibits the following
321 three characteristics.

[12] For approval of a biological license application, the safety of the proposed product must be demonstrated under section 351 of the Public Health Service Act (42 U.S.C. 262).

322
323 • The medical imaging agent meets *either* the safety-margin considerations *or* the clinical-
324 use considerations described below (see sections B.1 and B.2, respectively).

325 • The medical imaging agent is not a biological product[13, 14]

326 • The medical imaging agent does not predominantly emit alpha or beta particles

327 Note that under the safety margin criteria (see section IV.B), medical imaging agents that are
328 administered in low mass doses to humans (e.g., diagnostic radiopharmaceuticals) usually are
329 more likely to be considered Group 1 than those administered in higher mass doses.[15] There
330 are important exceptions, including cases where the medical imaging agents are likely to be
331 immunogenic (e.g., biological products) when the pharmacologic response exists at a low
332 mass dose, or when the medical imaging agents cause adverse reactions that are not dose-
333 related (e.g., idiosyncratic drug reactions).

334 We recommend that standard clinical safety evaluations be performed in all clinical
335 investigations of medical imaging agents, but we suggest that, for Group 1 agents, reduced
336 human safety monitoring may be appropriate in subsequent human trials.
337
338 • For example, human safety monitoring may be limited to recording adverse events and
339 monitoring only particular organs or tissues of interest for toxicity (such as organs that
340 showed toxicity in the animal studies, or the organs and tissues in which the medical
341 imaging agent localizes, which usually would include the liver and kidneys).
342
343 Persons having questions about whether a medical imaging agent is a Group 1 agent are
344 encouraged to contact FDA to discuss. Whether a medical imaging agent should be considered a
345 Group 1 or Group 2 agent may change during the course of a product's development. For
346 example, even if an agent is initially thought to be Group 1, the subsequent identification of
347 safety concerns could be reason to treat that agent as a Group 2 agent for the remainder of the
348 product's development.
349
350 *2. Group 2 Medical Imaging Agents*
351
352 For purposes of this guidance, Group 2 medical imaging agents are generally medical imaging
353 drugs or biological products that do not fall under the considerations for Group 1 medical
354 imaging agents. All biological products are assumed to be Group 2 agents unless the sponsor
355 demonstrates that its product lacks immunogenicity. Medical imaging agents that are

---

[13] Biological medical imaging products (e.g.., radiolabeled cells, monoclonal antibodies, monoclonal antibody fragments; see 21 CFR 600.3(h) for definition of a biological product) have the potential to elicit an immunogenic response. Because the development of antibodies following repeat or intermittent administration can alter the safety, pharmacokinetics, and biodistribution of such agents, we regard biological medical imaging products as Group 2 agents.

[14] See also the final regulation Adverse Experience Reporting Requirements for Licensed Biological Products (59 FR 54042; October 27, 1994).

[15] For example, the approved PET drug products meet the Group 1 criteria.

356 biologically active in animal studies or in human studies when administered at dosages that are
357 similar to those intended for clinical use should also be considered Group 2 agents.[16]
358
359 For Group 2 medical imaging agents, *standard clinical safety evaluations* should include serial
360 assessments of patient symptoms, physical signs, clinical laboratory tests (e.g., blood chemistry,
361 hematology, coagulation profiles, urinalyses), other tests (e.g., electrocardiograms as
362 appropriate), and adverse events. We recommend that additional specialized evaluations be
363 performed when appropriate (e.g., immunological evaluations, creatine kinase isoenzymes), or if
364 a particular toxicity is deemed possible based on animal studies or the known chemical or
365 pharmacological properties of the medical imaging agent. Although the extent of clinical
366 monitoring cannot be predetermined, we recommend that it be of sufficient duration to identify
367 possible effects that may lag behind those predicted by pharmacokinetic analyses. If some of
368 these standard clinical safety evaluations are felt to be unnecessary, this should be discussed with
369 the reviewing division. We recommend that sponsors seek FDA comment on the clinical safety
370 monitoring plans in clinical studies before such studies are initiated.
371
372 **B. Considerations For Groups 1 or 2**
373
374 *1. Safety-Margin Considerations*
375
376 Under the safety-margin considerations, medical imaging agents can be considered Group 1 if
377 the results of nonclinical studies ***and*** initial human experience are consistent with the conditions
378 outlined below:
379
380 a. Results of nonclinical studies
381
382 To be considered a Group 1 agent under the safety-margin considerations, we
383 recommend that a medical imaging agent have an adequately documented margin
384 of safety as assessed in the nonclinical studies outlined in the following list.[17]
385
386 • We recommend that the no-observed-adverse-effect level (NOAEL)[18] in
387 expanded-acute, single-dose toxicity studies in suitable animal species be at
388 least one hundred times (100x) greater than the maximal mass dose to be used
389 in human studies. We further recommend that such expanded, acute, single-
390 dose toxicity studies be completed before the medical imaging agent is
391 introduced into humans (see section III.B.1).

[16] Group 2 diagnostic radiopharmaceuticals can also include radionuclides and carriers that are known to be biologically active. This group includes radionuclides and carriers used at radiation doses or mass dosages that are higher than those used previously, including radionuclides and carriers that have been documented to produce adverse reactions.

[17] In addition, the medical imaging agent should meet the conditions described for the results of initial human experience (see section IV.B.1.b).

[18] For purposes of Groups 1 and 2 in this section of this guidance, the term *no-observed-adverse-effect-level (NOAEL)* is defined as the highest mass dose tested in animals with no adverse effects. (See guidance *A Harmonized Approach to Estimating the Safe Starting Dose for Clinical Trials of Therapeutics in Healthy Volunteers.*)

392
393 • We recommend that the NOAEL in safety pharmacology studies in suitable
394 animal species be at least one hundred times (100x) greater than the maximal
395 mass dose to be used in human studies. We further recommend that such
396 safety pharmacology studies be completed before the medical imaging agent
397 is introduced into humans (see section III.B.1).
398
399 • We recommend that the NOAEL in short-term, repeat-dose toxicity studies in
400 suitable animal species be at least twenty-five times (25x) greater than the
401 maximal mass dose to be used in human studies.[19] Short-term, repeat-dose
402 toxicity studies are conducted to evaluate the effects of exaggerated dose
403 regimens. Such regimens can reveal effects not detected in studies of small
404 numbers of patients, suggest effects to be monitored in clinical studies, and
405 reveal effects that might occur in sensitive individuals. Short-term, repeat-
406 dose toxicity studies can be performed either before the medical imaging
407 agent is introduced into humans, or concurrently with early human studies, but
408 we recommend that they be completed before phase 2 (see section III.B.1).
409
410 To establish these margins of safety, we recommend that the NOAELs be
411 assessed in properly designed and conducted studies and be appropriately
412 adjusted. *Appropriately adjusted* means that mass dose comparisons between
413 animals and humans should be suitably modified for factors such as body size
414 (e.g., body surface area) and otherwise adjusted for possible pharmacokinetic and
415 toxicokinetic differences between animals and humans (e.g., differences in
416 absorption for products that are administered orally).[20]
417
418 We recommend that Group 1 medical imaging agents also undergo other
419 nonclinical toxicological studies as described in section III.B.1, such as
420 genotoxicity, reproductive toxicity, irritancy studies, and drug-drug interaction
421 studies. See section III.B.1 for details and timing sequence.
422
423 i. Additional considerations
424
425 FDA may still consider a medical imaging agent Group 1 even if its NOAELs are
426 slightly less than the multiples specified above. For example, FDA will also take
427 into consideration, among other things, how close the NOAELs are to the
428 multiples specified above, the amount of safety information known about
429 chemically similar and pharmacologically related medical imaging agents, the

[19] Short-term, repeated-dose toxicity studies may identify toxicities associated with accumulation of a medical imaging agent or its metabolites. In addition, even if such accumulation is not anticipated (e.g., non-metabolized medical imaging agents with short half-lives), short-term repeated-dose toxicity studies may identify toxicities caused by repeated toxic insults, each of which may be below the threshold of detection in expanded-acute, single-dose toxicity studies.

[20] For example, if drug elimination is based on a physiologic function that reflects blood flow, we then recommend that scaling on body surface area be used.

nature of observed animal toxicities, and whether adverse events have occurred during initial human experience, including the nature of such adverse events (see section IV.B.1.b).

ii. Formulations used in nonclinical studies

We recommend that the formulation used to establish safety margins in nonclinical studies be identical to the formulation that will be used in clinical trials and that is intended for marketing. We also recommend that any differences in the formulations used in the clinical trials and nonclinical studies be specified so that any effect on the adequacy of the nonclinical studies can be determined. Bridging studies may be helpful when changes in the formulation are apt to change the pharmacokinetics, the pharmacodynamics, or safety characteristics of the drug.[21]

In some cases, it may be infeasible or impractical to administer the intended clinical formulation to animals in multiples of the maximal human mass dose specified above (e.g., the volume of such an animal mass dose may be excessive). We recommend that sponsors discuss their plans with FDA before studies are initiated. In these cases, alternative strategies can be employed, such as dividing the daily mass dose (e.g., into a morning and evening dose), or by using a more concentrated formulation of the medical imaging agent, or the maximal feasible daily mass dose can be administered.

b. Results of initial human experience

In addition to those considerations described above for nonclinical studies, FDA also intends to consider the following when evaluating whether a medical imaging agent is a Group 1 agent.

- Whether safety issues were identified during initial human use of the medical imaging agent in appropriately designed studies that include adequate and documented standard clinical safety evaluations. Identification of any adverse events during initial human use that were not predicted from effects observed in animals could be considered significant, regardless of severity. If adverse events occur at any time during human studies, we intend to conduct a risk assessment to determine whether the medical imaging agent should be reconsidered as a Group 2 medical imaging agent. This risk assessment will examine the type, frequency, severity, and potential attribution of the adverse events with respect to what is known about the pharmacology of the drug. For example, the safety profile of a specific drug class may be well known, so that the occurrence of a common, nonserious adverse event, such as headache, would not be of particular concern. However, in a drug class in which microparticles of varying sizes are administered, the occurrence of the same adverse event might be a signal of microcirculatory compromise.

[21] See guidance *S7A Safety pharmacology studies for human pharmaceuticals.*

- We recommend that human pharmacokinetic studies of the radiopharmaceutical be performed during phase 1 to collect information about the disposition of the radioactivity in humans. Such data help facilitate adequate comparisons of exposure between humans and the species used in the nonclinical studies and allow a more meaningful assessment of the relevance of the animal safety data (e.g., toxicokinetics).

2. Clinical Use Considerations

Another way to be considered a Group 1 agent is by adequately documenting extensive prior clinical use without development of a safety signal. This means showing that there were no human toxicity or adverse events with clinical mass doses (and activities, if applicable) of the agent, under conditions of adequate safety monitoring, and that the lack of human toxicity was adequately documented. We recommend that the methods used to monitor for adverse events be documented. Literature may be of limited value in establishing the clinical safety of a drug because most published studies focus on efficacy, with little or no description of any safety assessments.

An agent can be identified as Group 1 based on the clinical-use considerations at any time during drug development (e.g., after the conditions specified in this section have all been met).

C. Radiation Safety Assessment for All Diagnostic Radiopharmaceuticals

1. General Considerations

We recommend that an IND sponsor submit sufficient data from animal or human studies to allow a reasonable calculation of the radiation absorbed dose to the whole body and to critical organs upon administration to a human subject (21 CFR 312.23(a)(10)(ii)). At a minimum, we recommend that radiation absorbed dose estimates be provided for all organs and tissues in the standardized anthropomorphic phantoms established in the literature (e.g., by the Medical Internal Radiation Dose (MIRD) Committee of the Society of Nuclear Medicine). For diagnostic radiopharmaceuticals, we also recommend calculation of the *effective dose* as defined by the International Commission on Radiological Protection (ICRP) in its ICRP Publication 60 (this quantity is not meaningful for therapeutic radiopharmaceuticals).

When a diagnostic radiopharmaceutical is being developed for pediatric use, the radiation absorbed dose should be provided for all age groups in which the agent is intended to be used, as provided by standard anthropomorphic phantoms established in the literature (i.e., newborn, 1-year-old, 5-year-old, 10-year-old, and 15-year-old).

We recommend that the amount of the radiation absorbed dose delivered by internal administration of diagnostic radiopharmaceuticals be calculated by standardized methods, such as the absorbed fraction method described by the MIRD Committee and the ICRP.

We also recommend that the methodology used to assess radiation safety be specified including reference to the body models that were used. We recommend that the mathematical equations used to derive the time activity curves and the radiation absorbed dose estimates be provided along with a full description of assumptions that were made. We further recommend that sample calculations and all pertinent assumptions be listed and submitted. We recommend that the reference to the body, organ, or tissue model used in the dosimetry calculations be specified, particularly for new models being tested. If a software program was used to calculate the radiation doses, we recommend that you provide (1) a full description of the code, including official name, version number, and computing platform; (2) a literature citation for the code; and (3) photocopies of the code's output, preferably showing all of the user input data and model choices.

We recommend that safety hazards for patients and health care workers during and after administration of the radiolabeled product be identified, evaluated, and managed appropriately.

2. Calculation of Radiation Absorbed Dose to the Target Organs or Tissues

For established radionuclides used with a diagnostic agent (e.g., Tc-99m, In-111), we recommend that the following items be determined based on the average patient as defined by the MIRD phantom:

- The tissue or organ in which a significant accumulation of radioactivity occurs (i.e., *source* organ)
- The amount of radioactivity that accumulates in these tissues, expressed as a percentage of the administered activity
- The times at which radioactivity accumulation was observed in these tissues. We recommend that observations be made at two or more times during each phase of radioactivity accumulation or clearance from the source regions. If there is rapid accumulation in a region and nonexponential clearance, two to three time points may be sufficient to characterize the kinetic behavior. If there are two phases of clearance, we recommend at least two points of observation during each phase to adequately characterize the biokinetics. A description of the kinetic behavior of the activity accumulation and clearance from these tissues. This is most typically shown as *biological half-times* for accumulation and clearance, although other representations may be used.
- The time-integral of activity for the accumulation of radiopharmaceutical in these source tissues or organs. For purposes of this guidance, this *time-integral* is defined as the "cumulated activity" or "residence time" by the MIRD Committee in various publications.
- A description of how this time-integral was calculated. This should be based primarily on the accumulation and kinetic behavior in the source organs. We recommend that you specify the method used to calculate the time-integrals (e.g.,

565 numerical integration, regression analysis, or compartment model analysis). We also
566 recommend that you provide a description of how the terminal portion of the time-
567 activity curve for a given source region was integrated (e.g., assuming only physical
568 decay after the last data point, some rate of biological elimination estimated by two or
569 more of the later data points, or a fitted function continued to infinite time).

570 • A description of how these time-integrals for source regions were combined with
571 dose conversion factors to calculate the radiation absorbed dose to all target regions.
572 If hand calculations were performed, we recommend that you specify the source of
573 dose conversion factors and provide copies of all calculations. If an electronic
574 spreadsheet was used, we recommend that you provide printouts and electronic copies
575 of the spreadsheets to verify the formulas used. If a computer program was used, we
576 recommend that you provide a complete description of the code and version number
577 as well as documentation of the code input and output.

578 For new radionuclides used with diagnostic agents, the same principles apply and we
579 recommend that you provide the same information. If you want guidance on these
580 calculations, we recommend that you consult the appropriate review division.

581 *3. Maximum Radiation Absorbed Dose*
582
583 We recommend that the amount of radioactive material administered to human subjects
584 be the smallest radiation absorbed dose practical to perform the procedure while
585 providing an adequate diagnostic examination for evaluation by the physician.
586
587 We recommend that calculations include the radiation absorbed dose contributions made
588 by all potential radionuclide contaminants that may be present in the product.
589
590 We recommend that you perform calculations to anticipate possible changes in dosimetry
591 that might occur in the presence of diseases in organs that are critical in metabolism or
592 excretion of the diagnostic radiopharmaceutical. For example, renal dysfunction may
593 cause significant, slow-clearing accumulation in one or both kidneys (and thus a high
594 dose to kidneys and adjacent tissues) and/or a larger fraction of the administered activity
595 to be cleared by the hepatobiliary system (or vice versa).
596
597 We recommend that possible changes in dosimetry resulting from patient-to-patient
598 variations in antigen or receptor mass be considered in dosimetry calculations. For
599 example, a large tumor mass may result in a larger-than-expected radiation absorbed dose
600 to a target organ from a diagnostic radiopharmaceutical that has specificity for a tumor
601 antigen. (For the purposes of dose calculation, a primary tumor, without metastases, can
602 be regarded as part of the organ in which it arises and its activity can be added to that of
603 the organ.)
604
605 We recommend that the mathematical equations used to derive the estimates of the
606 individual organ time activity curves and the radiation absorbed doses be provided along
607 with a full description of assumptions that were made. We recommend that sample
608 calculations and all pertinent assumptions be listed.
609

610 We recommend that calculations of radiation absorbed dose estimates be performed
611 assuming freshly labeled material (to account for the maximum amount of radioactivity)
612 as well as the maximum shelf life of the diagnostic radiopharmaceutical (to allow for the
613 upper limit of accumulation of radioactive decay contaminants). We recommend that
614 these calculations:
615
616 • Include radiation absorbed doses from x-ray procedures that are part of the study (i.e.,
617 would not have occurred but for the study). The possibility of follow-up studies
618 should be considered for inclusion in the dose calculations.
619 • Be expressed as milligray (mGy) per megabecquerel (MBq) and as rad per millicurie
620 (mCi) of the administered radiopharmaceutical
621 • Be expressed as mGy and rad for a typical administered quantity of the
622 radiopharmaceutical
623 • Be presented in a tabular format and include individual radiation absorbed doses for
624 the target tissues or organs and the organs listed above in section IV.D.1
625

GLOSSARY

Effective dose: The sum of the weighted equivalent doses in all the tissues and organs of the body, given by the expression E = ∑WTHT,R, where WT is the weighting. Effective dose, defined in 1990 by the International Commission on Radiological Protection, allows the conversion of the risk from partial body irradiations to those of whole body irradiations.

Mass dose: The mass or weight of the ligand or carrier, including the radionuclide, administered to the subject.

No observed adverse effect level (NOAEL): The highest radiation absorbed dose tested in an animal species without adverse effects detected.

No observed effect level (NOEL): The highest radiation absorbed dose tested in an animal species with no detected effects.

Radiation absorbed dose: The energy absorbed per unit mass. This is the fundamental dosimetric quantity in radiological protection. Its unit is the joule per kilogram, which is given the special name gray (Gy). The older quantity was the rad, where 1 Gy = 100 rads.

Repeat-dose toxicity study: A study that investigates the toxicities produced when a pharmaceutical is administered repeatedly during a given period of more than 24 hours. A repeat-dose toxicity study evaluates the effects of exaggerated dose regimens. Usually all animals in a repeat-dose toxicity study are terminated the day after the final dose; however, a recovery period may be included in the design to test the reversibility of effects. An interim sacrifice is sometimes included to detect effects that may occur after a few doses.

Safety pharmacology study: A study that investigates the potential undesirable pharmacodynamic effects of a substance on physiologic functions in relation to exposure levels.

Special toxicology study: A study conducted when something about the nature of the drug or how it is used raises a concern, or when previous nonclinical or clinical findings on the product or a related product have indicated special toxicological concerns. Examples include a local irritation study conducted to test the effects of potential misadministration or extravasation.

Standard/expanded acute toxicity study: A study that investigates toxicities produced by a pharmaceutical when it is administered in one dose. During a period not exceeding 24 hours, doses may be split due to large volumes or high concentrations. An expanded acute toxicity study includes more measures of toxicities than a standard acute toxicity study.

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 2: Clinical Indications

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 2: Clinical Indications

Additional copies of this Guidance are available from:

Division of Drug Information HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Rockville, MD 20857
(Phone 301-827-4573)
Internet: http://www.fda.gov/cder/guidance/index.htm.

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
Internet: http://www.fda.gov/cber/guidelines.htm.
Mail: the Voice Information System at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Table of Contents

I. INTRODUCTION **2**

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS **3**

- **A. Contrast Agents** **3**
- **B. Diagnostic Radiopharmaceuticals** **4**

III. INDICATIONS FOR MEDICAL IMAGING AGENTS **4**

- **A. Structure Delineation** **5**
  - *1. Locating and Outlining Normal Anatomic Structures* *6*
  - *2. Distinguishing Between Normal and Abnormal Anatomy* *6*
- **B. Disease or Pathology Detection or Assessment** **7**
- **C. Functional, Physiological, or Biochemical Assessment** **8**
- **D. Diagnostic or Therapeutic Patient Management** **8**
- **E. Multiple or Other Indications** **10**

IV. CLINICAL TRIALS TO DETERMINE EFFECTIVENESS OF MEDICAL IMAGING AGENTS **10**

- **A. General Considerations for Establishing Accuracy and Validity of a Test** **11**
- **B. Clinical Usefulness** **12**
- **C. Defined Clinical Settings** **14**

APPENDIX: MEDICAL LITERATURE **16**

GLOSSARY **18**

Guidance for Industry[1]

Developing Medical Imaging Drug and Biological Products
Part: 2 Clinical Indications

This guidance represents 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. An alternative approach may be used if such 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

This guidance is one of three guidances intended to assist developers of medical imaging drug and biological products (*medical imaging agents*) in planning and coordinating their clinical investigations and preparing and submitting investigational new drug applications (INDs), new drug applications (NDAs), biologics license applications (BLAs), abbreviated NDAs (ANDAs), and supplements to NDAs or BLAs. The three guidances are*: Part 1: Conducting Safety Assessments*; *Part 2: Clinical Indications*; and *Part 3: Design, Analysis, and Interpretation of Clinical Studies*.

Medical imaging agents generally are governed by the same regulations as other drugs or biological products. However, because medical imaging agents are used solely to diagnose and monitor diseases or conditions as opposed to treat them, development programs for medical imaging agents can be tailored to reflect these particular uses. Specifically, this guidance discusses our recommendations on selecting and studying clinical indications for medical imaging agents administered in vivo.[2]

---

[1] This guidance has been prepared by the Division of Medical Imaging and Radiopharmaceutical Drug Products and the Office of Therapeutics Research and Review in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration.

[2] In response to the requirements of the Food and Drug Administration Modernization Act of 1997, FDA amended the drug and biologics regulations (21 CFR 315 and 601) by adding provisions for the evaluation and approval of in vivo radiopharmaceuticals used in the diagnosis or monitoring of diseases (64 FR 26657, May 17, 1999). This guidance elaborates on the provisions contained in that regulation.

36
37 FDA's guidance documents, including this guidance, do not establish legally enforceable
38 responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should
39 be viewed only as recommendations, unless specific regulatory or statutory requirements are
40 cited. The use of the word *should* in Agency guidances means that something is suggested or
41 recommended, but not required.
42
43 A glossary of common terms used in diagnostic medical imaging is provided at the end of this
44 document.
45
46
47 **II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS**
48
49 This guidance discusses medical imaging agents that are administered in vivo and are used for
50 diagnosis or monitoring with a variety of modalities, such as radiography, computed tomography
51 (CT), ultrasonography, magnetic resonance imaging (MRI), and radionuclide imaging. The
52 guidance is not intended to apply to the development of in vitro diagnostic or therapeutic uses of
53 these agents.[3]
54
55 Medical imaging agents can be classified into at least two general categories, contrast agents and
56 diagnostic radiopharmaceuticals.
57
58 **A. Contrast Agents**
59
60 As used in this guidance, a contrast agent is a medical imaging agent used to improve the
61 visualization of tissues, organs, and physiologic processes by increasing the relative difference of
62 imaging signal intensities in adjacent regions of the body. Products include, but are not limited
63 to (1) iodinated compounds used in radiography and CT; (2) paramagnetic metallic ions (such as
64 ions of gadolinium, iron, and manganese) linked to a variety of molecules and used in MRI; and
65 (3) microbubbles, microaerosomes, and related microparticles used in diagnostic
66 ultrasonography.

---

[3] The guidance is not intended to apply to the development of research drugs that do not provide direct patient benefit with respect to diagnosis, therapy, prevention, or prognosis, or other clinically useful information. These include radioactive drugs for research that are used in accordance with 21 CFR 361.1. Section 361.1(a) states that radioactive drugs (defined in 21 CFR 310.3(n)) are generally recognized as safe and effective when administered under specified conditions to human research subjects in the course of a project intended to obtain basic information about the metabolism of a radioactively labeled drug or about human physiology, pathophysiology, or biochemistry. However, if a radioactive drug is used for immediate therapeutic, diagnostic, or similar purposes or to determine the safety and effectiveness of the drug in humans, or if the radioactive drug has a pharmacological effect in the human body, an IND is required. FDA is developing a guidance on determining when research with radioactive drugs may be conducted under § 361.1.

The Agency recognizes the potential of imaging agents as research tools for aiding the development of therapeutic drugs, and some of the principles in the guidance may be applicable to such research. Sponsors of such imaging research agents are urged to contact the Division of Medical Imaging and Radiopharmaceutical Drug Products for advice on development of the imaging research agent.

B. Diagnostic Radiopharmaceuticals

As used in this guidance, a *diagnostic radiopharmaceutical* is (1) an article that is intended for use in the diagnosis or monitoring of a disease or a manifestation of a disease in humans and that exhibits spontaneous disintegration of unstable nuclei with the emission of nuclear particles or photons or (2) any nonradioactive reagent kit or nuclide generator that is intended to be used in the preparation of such an article.[4] The FDA interprets this definition to include articles that exhibit spontaneous disintegration leading to the reconstruction of unstable nuclei and the subsequent emission of nuclear particles or photons.

Diagnostic radiopharmaceuticals are generally radioactive drugs or biological products that contain a radionuclide that typically is linked to a ligand or carrier.[5] These products are used with planar imaging, single photon emission computed tomography (SPECT), positron emission tomography (PET), or with other radiation detection probes.

Diagnostic radiopharmaceuticals used for imaging typically have two distinct components.

- A radionuclide that can be detected in vivo (e.g., technetium-99m, iodine-123, indium-111).

  The radionuclide typically is a radioactive molecule with a relatively short physical half-life that emits radioactive decay photons having sufficient energy to penetrate the tissue mass of the patient. These photons can then be detected with imaging devices or other detectors.

- A nonradioactive component to which the radionuclide is bound that delivers the radionuclide to specific areas within the body.

  This nonradionuclidic portion of the diagnostic radiopharmaceutical often is an organic molecule such as a carbohydrate, lipid, nucleic acid, peptide, small protein, or antibody.

As technology advances, new products may emerge that do not fit into the traditional categories of contrast agents and radiopharmaceuticals (e.g., agents for optical imaging, magnetic resonance spectroscopy, combined contrast and functional imaging). It is anticipated, however, that the general principles discussed here could apply to these new diagnostic products. Developers of these products should contact the appropriate reviewing division for advice on product development.

III. INDICATIONS FOR MEDICAL IMAGING AGENTS

[4] 21 CFR 315.2 and 601.31.

[5] In this guidance, the terms *ligand* and *carrier* refer to the entire nonradionuclidic portion of the diagnostic radiopharmaceutical.

106 The labeled indications for medical imaging agents fall within the following general categories:
107
108 • Structure delineation
109 • Disease or pathology detection or assessment
110 • Functional, physiological, or biochemical assessment
111 • Diagnostic or therapeutic patient management
112
113 The above categories do not represent a hierarchy or progression (e.g., a *structure delineation*
114 indication does not need to precede a *disease assessment* indication). In addition, indications
115 from different categories could be granted for the same imaging agent. Approval also may be
116 possible for categories of indications not listed above.
117
118 Under section 505(d) of the Federal Food, Drug, and Cosmetic Act (the Act) (21 U.S.C. 355(d)),
119 FDA cannot approve a new drug application (NDA) unless it contains adequate tests
120 demonstrating whether the proposed drug product is safe for use under the conditions prescribed,
121 recommended, or suggested in its proposed labeling.[6] All drugs have risks, including risks
122 related to the intrinsic properties of the drug, the administration process, the reactions of the
123 patient, and incorrect diagnostic information. Incorrect diagnostic information includes
124 inaccurate structural, functional, physiological, or biochemical information; false positive or
125 false negative diagnostic determinations; and information leading to inappropriate decisions in
126 diagnostic or therapeutic management. Even if risks are found to be small, all drug development
127 programs must obtain evidence of drug effectiveness under section 505 of the Act. Simply
128 generating an image, for which the implications to the patient are not understood, does not
129 confer benefits to the patient.
130
131 In determining the most appropriate indication for a medical imaging agent, special
132 considerations may apply to agents that may pose significant patient risk, for example, biological
133 medical imaging agents that are frequently immunogenic. The development of antibodies after
134 intermittent, repeated administration can alter the pharmacokinetics, biodistribution, safety,
135 and/or imaging properties of such agents and, potentially, of immunologically related agents. For
136 agents that pose significant risk and where the clinical benefit is generally not readily apparent,
137 an indication of *disease or pathology detection or assessment* or *diagnostic or therapeutic*
138 *patient management* is more appropriate. If one of the other indications (i.e*., structure*
139 *delineation* or *functional, physical or biochemical assessment*) will be sought for an agent that
140 may pose significant patient risk, we recommend that the development plan be discussed with
141 the review division.
142
143 **A. Structure Delineation**
144
145 As described in the following sub-sections, at least two types of labeled indications for structure
146 delineation are possible: (1) locating and outlining normal (or variants of normal) anatomic
147 structures and (2) distinguishing between normal and abnormal anatomy in a defined clinical
148 setting. Ordinarily, the ability to locate and outline normal structures or distinguish between

[6] For approval of a biological license application, the safety of the proposed product must be demonstrated under section 351(a) of the Public Health Service Act (42 U.S.C. 262(a)).

149 normal and abnormal anatomy can *speak for itself* with respect to the clinical value of the
150 information and will not require additional information substantiating clinical usefulness.
151
152 *1. Locating and Outlining Normal Anatomic Structures*
153
154 We recommend that a medical imaging agent intended for this type of indication be able
155 to locate and outline normal (or variants of normal) anatomic structures. We recommend
156 that the product clarify the spatial relationship of the visualized normal structure with
157 respect to other body parts or structures. Such a medical imaging agent could distinguish
158 normal structures that cannot be seen well with other imaging agents or modalities. For
159 example, a contrast agent developed to image the normal parathyroid glands could be
160 clinically useful because it could help surgeons plan and perform thyroid surgery.
161
162 *2. Distinguishing Between Normal and Abnormal Anatomy*
163
164 We recommend that a medical imaging agent intended for this type of indication be able
165 to locate and outline both normal and abnormal anatomic structures. We recommend that
166 the agent also clarify the spatial relationships of the normal and abnormal anatomic
167 structure(s) with respect to other body parts or structures. Imaging agents that identify
168 abnormalities common to one or more disease entities (and therefore not specific to a
169 particular disease) could be eligible for a structural indication. Examples of this type of
170 agent include:
171
172 • An agent that nonspecifically enhances the airway lumen to distinguish dilated
173 bronchi from normal bronchi and categorizes the bronchiectasis anatomically
174 (e.g., as cylindric, sacculated, or fusiform)
175 • An agent that nonspecifically enhances the joint cavity to evaluate and describe
176 meniscal or ligamentous injuries of the knee
177 • An agent that outlines the vascular system to identify structural narrowing,
178 dissections, aneurysms, and relationships to normal vasculature
179 • A contrast agent that localizes or outlines masses
180
181 In the preceding examples, the agent's ability to outline abnormal anatomy may also be
182 supportive of a disease detection indication in a specific population (section III.B). If the
183 sponsor can demonstrate that use of the agent provides clinical benefit in this population,
184 a disease detection indication might be appropriate.
185

186 **B. Disease or Pathology Detection or Assessment**
187
188 We recommend that a medical imaging agent intended for disease or pathology detection be able
189 to detect and locate a specific disease or pathological state in at least one defined clinical
190 setting.[7] The medical imaging agent could be used alone or in combination with other diagnostic
191 procedures to achieve this labeled indication.
192
193 Examples of medical imaging agents for which this type of indication may be appropriate
194 include:
195
196 • An agent that can bind to multiple regions of the brain but is intended to detect or assess
197 the extent of a specific neurological disease, such as Parkinson's disease
198 • A radiolabeled monoclonal antibody that can attach to a unique tumor antigen to detect
199 the presence of, or extent of, a mass with this tumor antigen (e.g., breast cancer)
200
201 We recommend that efficacy trials for these indications be conducted in subjects presenting for
202 diagnostic evaluation of a specific disease or condition in a defined clinical setting. This is
203 because the likelihood of disease or the spectrum of disease (e.g., severity or stage) is dependent
204 on the clinical setting. Examples of two common clinical settings are (1) providing a diagnosis
205 in patients with suspected disease and (2) monitoring and assessing the extent, rate of
206 progression, or other aspects of the specific disease in patients previously diagnosed with the
207 disease. An indication of detection of disease or pathology in an asymptomatic population (a
208 screening indication) may be appropriate if the sensitivity of the imaging modality is high
209 enough and the rate of false positives is low enough (see also diagnostic or therapeutic patient
210 management, section III.D).
211
212 It is likely that the clinical usefulness and the diagnostic performance of the medical imaging
213 agent will differ in each clinical setting.[8] We recommend that if a medical imaging agent is
214 being developed to diagnose a particular disease, efficacy trials generally enroll subjects in
215 whom the disease status is unknown, but in whom specific aspects of the clinical presentation
216 have led to the desire for more diagnostic information. That is, we recommend that the trials
217 include the intended population in the appropriate clinical setting. Data from subjects known
218 definitely to have (or to not have) the disease of interest may be of limited value because
219 estimates of diagnostic performance derived from a known disease population may not apply to
220 performance in the intended population.
221
222
223

[7] See section IV.C for a definition of *defined clinical setting*.

[8] Studying patients with known disease provides information useful in developing a hypothesis for testing in subsequent clinical trials. Typically, such clinical settings are not used to establish efficacy in disease or pathology detection.

C. Functional, Physiological, or Biochemical Assessment

We recommend that a medical imaging agent intended to provide functional, physiological, or biochemical assessment be able to evaluate the function, physiology, or biochemistry of a tissue, organ system, or body region. This type of indication could apply to agents used to detect either a reduction or an increase of a normal functional, physiological, or biochemical process. The indication *functional, physiological, or biochemical assessment* could be limited to assessment of functional, physiological, or biochemical processes when disturbances of these processes are common to several diseases or conditions and they are not diagnostic for any particular disease or condition.

The indication *functional, physiological, or biochemical assessment* is appropriate for patients when evaluations of functional, physiological, or biochemical aspects of a tissue, organ, or body region would provide clinically useful information.

Examples of medical imaging agents with *functional, physiological, or biochemical assessment* indications include:

- A contrast agent to assess cardiac ejection fraction or myocardial wall motion
- A radiopharmaceutical that assesses metabolism of a substrate where the normal pattern of metabolism in that organ or tissue is well known

To establish efficacy in clinical studies, we recommend that the functional, physiological, or biochemical measurements of the medical imaging agent be compared with those of a reference product or a procedure of known high validity (i.e., a truth standard). Ideally, we recommend that the high validity of this reference product or truth standard be documented thoroughly and critically before its use in clinical studies intended to demonstrate effectiveness of the test-imaging agent. We recommend that a functional indication be studied in the wide spectrum of diseases and disease severity states that affect the functional endpoint. For example, a sponsor might seek an indication of measuring myocardial left ventricular function. To ensure that the test is valid in the patient population most likely to be referred for testing, the sponsor might design studies that include subjects with different cardiac diseases, such as dilated cardiomyopathy, valvulopathy, hypertrophic cardiomyopathy, and myocardial infarction, including subjects with normal function as well as those with mild, moderate, and severe dysfunction. In that case, separate studies for each disease would not be needed.

If no standard of truth applies to the proposed use of a medical imaging agent for functional, physiological, or biochemical assessment, we recommend that a clinical trial be conducted to determine that the findings are clinically useful (see section IV.B).

D. Diagnostic or Therapeutic Patient Management

We recommend that a medical imaging agent intended for the indication *diagnostic or therapeutic patient management* be able to improve patient management decisions (e.g., the need for further diagnostic testing or the use of specific therapeutic interventions) or improve patient

269 outcomes when used in a defined clinical setting.[9] Included in this indication is the ability to
270 provide information (such as the presence of a certain receptor in a type of cancer patient) that
271 can predict survival or patient response to a particular type of therapeutic drug.
272
273 To obtain approval for a diagnostic or therapeutic patient management indication, we
274 recommend that adequate and well-controlled investigations demonstrate that patient
275 management decisions or outcomes are, in fact, improved by use of the medical imaging agent.
276 The medical imaging agent can be used alone or in combination with other diagnostic procedures
277 to achieve this labeled indication. Studies might involve randomization into one arm with
278 testing and patient management based on the testing results in accordance with a prespecified
279 algorithm compared to a nontesting arm that proceeded to treatment as defined by current
280 standards. Patient outcomes such as recovery, survival, and response rates that are based on the
281 treatment standard could be collected and compared.
282
283 Examples of medical imaging agents for which this type of indication may be appropriate are:
284
285 • Products shown to provide improved clinical decisions about whether suspected cardiac
286 patients should undergo further invasive, diagnostic testing, such as with coronary
287 angiography (i.e., use for diagnostic patient management)
288 • Products that predict whether a patient has a better prognosis with tumor resection
289 instead of with chemotherapy (i.e., use for therapeutic patient management)
290
291 We recommend that the trials demonstrate that diagnostic or therapeutic management is
292 improved when using the medical imaging agent compared to management without use of the
293 medical imaging agent. The medical imaging agent can be used in conjunction with other tests
294 to influence a patient diagnostic or therapeutic management decision. We suggest that it would
295 not be sufficient simply to demonstrate that the results of the test drug were used to direct a
296 change in patient management, even to an intervention that is well established. Rather, we
297 recommend that the sponsor establish whether the change was better or worse for the patient.
298 For example, when using a new imaging agent in determining whether to perform breast biopsy
299 versus repeat clinical breast examination followed by mammography, the sponsor should show
300 whether use of the test drug results in fewer or greater numbers of unnecessary biopsies or
301 undiagnosed cancers than use of a comparator. We recommend that this principle also be
302 applied to studies to demonstrate a therapeutic patient management claim: the sponsor should
303 show that use of the new imaging agent leads to better patient therapy choices than result from
304 the use of existing methods of managing therapy. If sponsors do not wish to perform such
305 follow up, we recommend that they instead seek an indication for disease or pathology detection
306 or assessment.
307
308 To obtain the indication *diagnostic or therapeutic patient management* for a medical imaging
309 agent that identifies unrecognized disease in asymptomatic individuals (e.g., used in a screening
310 setting), we recommend that a sponsor show that use of the test decreases morbidity or mortality,

[9] See section IV.C for a definition of *defined clinical setting*.

311 or provide existing data that show that early detection and treatment of the disease decreases
312 morbidity or mortality.
313
314 **E. Multiple or Other Indications**
315
316 The indication categories outlined above are flexible, and indications for medical imaging agents
317 need not be mutually exclusive. A labeled indication can include several indication categories.
318 For example, a diagnostic radiopharmaceutical could be developed as an aid in the diagnosis of
319 lung cancer for the labeled indication *disease or pathology detection or assessment*. This
320 diagnostic radiopharmaceutical could also be evaluated in subpopulations of patients with lung
321 cancer for its ability to provide information that leads directly to appropriate therapeutic
322 management decisions (e.g., using test results to determine what combination of surgery,
323 radiotherapy, and chemotherapy is most appropriate).
324
325 Structural and functional aspects of diseases or conditions sometimes are evaluated together with
326 imaging in clinical practice (e.g., use of a contrast agent to evaluate cardiac anatomy and
327 segmental wall motion). In such cases, we recommend that clinical studies evaluate the effect of
328 the imaging agent on assessments of both structure and function.
329
330 Functional evaluations of diseases or conditions may be accomplished for various purposes. For
331 example, a drug may have a functional indication for the evaluation of cardiac ejection fraction.
332 Subsequently, the drug may be developed for a therapeutic management indication for the
333 evaluation of perfusion or wall motion abnormalities to predict response to surgical intervention.
334
335 For indications that do not fall within the categories identified above (e.g., providing prognostic
336 information based on imaged gene expression), we recommend that the applicant or sponsor
337 consult FDA on the nature of the desired labeled indication and how to establish effectiveness
338 for it.
339
340
341 **IV. CLINICAL TRIALS TO DETERMINE EFFECTIVENESS OF MEDICAL**
342 **IMAGING AGENTS**
343
344 In general, establishing effectiveness has two components: (1) establishing the accuracy of the
345 test and (2) establishing the clinical value of the test. In some cases, a test that provides accurate
346 information in describing a clinical condition is of well-established value. Generally, this is true
347 for proposed indications for structure delineation and disease or pathology detection or
348 assessment. When there are established methods of seeking similar information and the only
349 issue is comparing the accuracy of the new and old methods, the clinical usefulness of the
350 indication need not be reestablished. Many functional, physiological, or biochemical
351 assessments are similarly well established as useful (ejection fraction, renal function, myocardial
352 wall motion) but others (glucose utilization by various parts of the body, presence of serotonin
353 receptors, cerebral blood flow, palmitate metabolism) may not be. Where the clinical value of
354 valid information is not established, we recommend that additional information establishing its
355 value be developed. This recommendation applies to all drugs, including therapeutic drugs, for
356 which the indication or mechanism of action for an indication is not accepted or well understood

357 in the medical community.
358
359 Demonstration of improved patient management means more than assessment of the accuracy of
360 the test. Either by reference to prior data or through new trials, we recommend that this claim
361 show that the test really makes a difference in outcome or management. Of course, the impact of
362 a test may be considered obvious (e.g., staging of breast cancer is disease or pathology detection
363 or assessment indication). In some cases, the test will have plain therapeutic implications, as
364 would be the case for effective staging of some other malignancies, although, in many cases, trial
365 data should be collected.
366
367 We recommend that investigations establish the validity (generally assessed by describing the
368 sensitivity, specificity, positive predictable value and negative predictive value in relevant
369 settings) and reliability (how reproducible the test results are) of the imaging agent. These test
370 characteristics can provide information on risk-benefit as well, including estimates of risk of
371 incorrect diagnosis. Safety information obtained in studies (see the companion guidance *Part 3:*
372 *Design, Analysis, and Interpretation of Clinical Studies*) will also contribute to an Agency risk-
373 benefit assessment. The clinical usefulness of an imaging agent may be obvious from a
374 description of what it can demonstrate (or supportable by evaluation of the literature), or it may
375 be appropriate to demonstrate the agent's usefulness. We recommend that clinical studies and
376 related methods for establishing effectiveness be performed in defined clinical settings that
377 reflect the proposed indications.
378

379 **A. General Considerations for Establishing Accuracy and Validity of a Test**

380
381 To establish efficacy in clinical studies, we recommend that the accuracy and/or validity of the
382 structural delineation, functional, physiological, or biochemical assessment and disease or
383 pathology detection generally be demonstrated by comparing the performance of the medical
384 imaging agent with that of a reference product or a truth standard in a relevant clinical setting.
385
386 To provide adequate estimates of the validity and reliability of the medical imaging agent over
387 the full range of conditions for which it is intended to be used, we recommend that medical
388 imaging agents be evaluated in studies with appropriate representation of sufficient numbers of
389 subjects (1) with and without the abnormalities or diseases in question (over the full spectrum of
390 the condition or disease presentation) and (2) with other conditions, processes, or diseases that
391 could affect the interpretation of the imaging results (e.g., inflammation, neoplasm, infection,
392 trauma). We recommend that sponsors justify the inclusion or exclusion of selected
393 subpopulations during clinical development. We recommend that studies of agents for
394 functional, physiological, or biochemical assessment indications provide a quantitative or
395 qualitative understanding of how the measurement varies in normal and abnormal subjects or
396 tissues, including the variable's normal range, distribution, and confidence intervals in these
397 subjects or tissues. We believe it is critical to identify the range that is normal and the values
398 that indicate an abnormality. When possible, we recommend that the minimum detectable limits
399 and reproducibility of the measurement be assessed.
400

401 Reproducibility assessments are most meaningful when performed within the same subject.
402 However, under some circumstances this practice might be unethical, in which case the sponsor
403 should consider alternative approaches to testing reproducibility (e.g., in animals).
404
405 In cases when a valid reference product or a truth standard is unavailable or infeasible, the
406 validity of the information obtained can be demonstrated in clinical studies showing a beneficial
407 clinical outcome. We recommend that the sponsor discuss these issues with the Agency prior to
408 initiation of phase 3 studies.
409
410 **B. Clinical Usefulness**
411
412 Under section 505 of the Act (or, for a biologic, section 351 of the PHS Act) and its
413 implementing regulations, FDA cannot approve a drug without evidence that the drug's benefit
414 to patients outweighs its risks:
415
416 • Sponsors are not asked to re-demonstrate the benefits already shown by diagnosing a
417 patient's specific disease or defining anatomy or functional status. However, under
418 section 505 of the Act or section 351 of the PHS Act, the sponsor must demonstrate that
419 the agent's benefits justify the risks.
420 • For an indication for which the benefits of an imaging agent have not yet been shown,
421 such as imaging a new biochemical process, a sponsor should conduct clinical trials to
422 demonstrate the agent's clinical prognostic value, the diagnostic performance
423 characteristics of the test agent compared to existing testing, the test's ability to predict
424 appropriate therapy, or the test's ability to help select appropriate further diagnostic
425 testing over existing testing.
426
427 The use of medical imaging agents without defined benefits and without an understanding of
428 how the imaging results can be used for patient management might cause harm to patients even
429 if the agent has low toxicity. Such harm might include (1) conducting unnecessary diagnostic
430 testing based on the results of the agent, (2) directing patients to invasive procedures or
431 inappropriate or unnecessary therapy, and (3) creating unnecessary patient anxiety from
432 *abnormal* test results.
433
434 Medical imaging results may have clinical usefulness in some settings but not in others; it is,
435 therefore, important to prospectively define and study the imaging agent in the clinical setting of
436 intended use. We recommend that a medical imaging agent be able to provide accurate and
437 reliable information that, in one of a number of ways, facilitates clinical management, including
438 (1) helping make an accurate diagnosis, (2) contributing to beneficial clinical outcome (e.g., by
439 helping choose the right therapy), or (3) providing accurate prognostic information. All
440 indications under section III should reflect these benefits, which are then weighed by FDA
441 against the agent's risks as part of an approval decision. Once clinical usefulness is established,
442 other benefits of imaging agents, such as safety advantages and enhanced convenience to
443 patients over existing marketed products, can be considered.
444

445 Depending on the specific indication, clinical usefulness can generally be established in two
446 ways: (1) by direct demonstration in studies carried out during clinical development and (2) by
447 reference to historical data. In circumstances when the measure is well established as useful in
448 the medical literature, the clinical benefit of the measure does not need to be re-established (e.g.,
449 ejection fraction or myocardial wall motion are widely used measures of cardiac function with
450 known prognostic and therapeutic implications). Even if the new measure has not been used
451 before, we suggest that clinical usefulness can be established historically when the information
452 being obtained has been shown to be useful when obtained by other means. For example, if a
453 product is able to establish the early detection of colon polyps without the need for colonoscopy,
454 the clinical benefit of the use of this product can be inferred because treatment is available for
455 this disease (polypectomy), and the test would allow people to avoid unnecessary colonoscopy
456 (i.e., clinical usefulness has been established indirectly). In such situations, clinical usefulness
457 can be documented by a critical and thorough analysis of the medical literature and any historical
458 precedents.
459
460 For indications in which it cannot be established from prior knowledge, we recommend that
461 clinical usefulness be established through new trials during development. For example, we
462 recommend that clinical usefulness be established directly for a medical imaging agent that has
463 been shown in a research setting to bind specifically to particular receptors, but where it has not
464 yet been established that assessment of such binding adds to the accuracy of diagnosis,
465 contributes to beneficial clinical outcome, or provides accurate prognostic information. For
466 novel technologies relying on mechanisms for imaging never approved before, we recommend
467 that a plan for establishing clinical usefulness be incorporated into the development plan of a
468 medical imaging agent. In general, we recommend that clinical usefulness be evaluated
469 prospectively in the principal clinical studies of efficacy. We recommend that sponsors assess
470 how the novel technology imaging results are used and how usefulness to the patient is
471 confirmed.
472
473 For a contrast agent to be considered clinically useful, we recommend that, when used in
474 combination with an imaging device, the agent be able to provide useful information or other
475 advantages (such as improved imaging time or convenience) beyond that obtained by the
476 imaging device alone. That is, we recommend that imaging with the contrast agent have added
477 benefit when compared to imaging without the contrast agent.
478
479 To illustrate how effectiveness could be evaluated, consider the following possible approaches:
480
481 1. Compare the new test and the established (comparator) test, which could be either
482 another test or a truth standard, such as pathology. Ideally, both a comparator and
483 a truth standard are employed so that the diagnostic performance measures of the
484 new test can be compared to those of the comparator. We recommend that the
485 population studied include the spectrum of presenting patients that would be
486 expected to undergo the new test, and that standard analyses be performed on
487 sensitivity and specificity, positive predictive value, and negative predicative
488 value. If the comparator test is established as clinically useful through controlled
489 clinical trials or literature and is considered the standard of care by the practicing

community, comparing the new test to the comparator could be sufficient to demonstrate the usefulness of the new test.

2. Compare the new test added to the current standard test battery to a truth standard (such as pathology). If the new test added to the standard battery shows greater sensitivity and specificity than the standard test battery without the new test, that result alone could be sufficient. Or, if the new test detects some lesions that the standard tests miss (greater sensitivity) and the result is of accepted clinical value (i.e., leads to improved patient management and has a very low false positive rate), that result alone could be sufficient.

Note: Situation 2 would be similar to imaging with and without the new test to determine the contribution of the new imaging test.

For 1 and 2, the results of the new test would be presumed to be of prognostic, therapeutic, or diagnostic value, and the new imaging drug would be presumed to improve these aspects. If that is not the case, we recommend that the value be documented through a randomized clinical trial. The new test (or the new test added to the standard testing battery) could be compared to standard testing without the new test to determine if the new test improves clinical outcomes or prognosis. Refer to Part 3 of the medical imaging guidances, section IV.D.1, for additional discussion.

C. Defined Clinical Settings

We recommend that a *defined clinical setting* reflect the circumstances and conditions under which the medical imaging agent is intended to be used.[10] Generally, the choice of anticipated labeled indications will determine the clinical setting for the trials. In some cases, an appropriately designed trial may contain several clinical settings.

For example, a medical imaging agent intended to detect prostate cancer (a disease specific indication) could be developed for use in different defined clinical settings such as:

- For asymptomatic, healthy men for early detection screening
- For use in men presenting with a high clinical index of suspicion for prostate cancer either by physical examination or abnormal prostate specific antigen testing
- For use in men with existing diagnosis of prostate cancer to evaluate recurrence

[10] Note that use of a *defined clinical setting* in studies of medical imaging agents also tends to anchor both the *pretest probability* and the *spectrum* (e.g., severity or stage) of the disease or condition under study. Thus, when evaluated in a defined clinical setting, diagnostic performance measures that vary with the pretest probability of the disease or condition (e.g., positive and negative predictive values, accuracy), or that can vary with the spectrum of the disease or condition (e.g., sensitivity, specificity, positive and negative predictive values, accuracy) tend to take on values that are relatively constant for that defined clinical setting. See section III B.

526 We recommend that the circumstances and conditions under which the medical imaging agent is
527 intended for use be evaluated in clinical trials and be described in the labeling using the
528 following mechanisms:
529
530 1. Specify aspects of the medical history and physical examination that are pertinent for
531 determining the likelihood of the disease or condition that is in question. For example:
532
533 • A medical imaging agent intended to detect breast cancer might be evaluated for use
534 in the assessment of (1) otherwise healthy women over 40 years of age, (2) women
535 with a family history of breast cancer, or (3) women presenting with palpable breast
536 masses or abnormal mammograms.
537
538 2. Specify a patient population that is at a particular step in the diagnostic or patient
539 surveillance sequence. For example:
540
541 • A diagnostic radiopharmaceutical may be intended to evaluate patients in an
542 emergency room with equivocal clinical and laboratory findings of a myocardial
543 infarction, or to evaluate the location and extent of a myocardial infarction in patients
544 with definitive findings.
545 • An agent may be intended for use in an outpatient office setting to monitor disease
546 progression. Typically, we recommend that such an imaging agent be studied using
547 repeated, periodic surveillance imaging of ambulatory patients in an outpatient
548 setting.
549
550 3. Specify any other diagnostic assessments that are to be performed in the evaluation of
551 this patient population. We recommend that this delineation include describing how the
552 medical imaging agent should be used with respect to other diagnostic tests or
553 evaluations, including (1) whether the medical imaging agent is intended to be used
554 together with, or as a replacement for, other diagnostic tests or modalities and (2) how
555 the use of the medical imaging agent is influenced by the results of other diagnostic
556 evaluations.
557
558 Pooling efficacy data (additive derivation of summary statistics) across defined clinical settings
559 may only be of limited value because differences in disease prevalence and in pathophysiology
560 may result in different diagnostic performance (sensitivity, specificity, positive and negative
561 predictive value) in different settings. Pooled results may suggest that the product is useful in all
562 the evaluated clinical settings and may obscure the evidence of differential usefulness in each
563 one of the settings. Of course, data from independent trials in different clinical settings may be
564 useful in determining the overall labeling in one or more clinical settings. The number and type
565 of populations to be studied depends on the type of the indication and clinical uses sought by the
566 sponsor. If there are data showing that the benefits from use of a medical imaging agent in a
567 particular clinical setting exceed the risks, that can be reflected on the labeling.
568
569
570

571
572 **APPENDIX: MEDICAL LITERATURE**
573
574 On occasion, the medical literature may provide critical information on various aspects of the
575 safety or efficacy of a product. Considerations in the use of the medical literature are described
576 in the FDA guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug*
577 *and Biological Products*. In applying these fundamental principles to imaging trials, we
578 recommend that sponsors consider whether the methods section in a relevant literature article
579 describes a prospective protocol in sufficient detail to assess the strengths and weaknesses of the
580 protocol design features discussed in this medical imaging guidance. For example, the design
581 features that are critical include selection of the patient population, clinical setting, image
582 handling, image reading plan for the test product and the standard of truth, use of an accepted
583 standard of truth, the statistical plan, and use of appropriate steps to eliminate bias. As literature
584 studies are often completed for purposes other than drug approval, the relevance of the selected
585 endpoints to the proposed indication should be justified.
586
587 We recommend that a critical review of the literature present the method used for the literature
588 search, the criteria used to review the data, and the criteria used to determine the applicability of
589 the results. Although we recommend that each article be reviewed and summarized, we also
590 recommend that the key articles be discussed more extensively.
591
592 Typically, articles in the imaging literature provide limited data on safety, so that additional
593 safety studies may be called for. Other information that can be supplied either fully or partially
594 by the literature include:
595
596 • Information on human drug safety: population exposed, types of adverse events and how they
597 were monitored and reported, reliability of data collection
598 • Pharmacology
599 • Toxicology studies
600 • Biopharmaceutical information
601
602 In the FDA guidance for industry, *Providing Clinical Evidence of Effectiveness for Human Drug*
603 *and Biological Products*, the Agency discusses the use of the medical literature by sponsors. We
604 recommend:
605
606 • Independent substantiation of experimental results. Multiple studies from different authors
607 provide greater support.
608 —Replication of findings in usually two or more adequate and well-controlled human
609 investigations
610 —Similar study questions, populations, diseases or conditions or indications being
611 studied with the same imaging agent
612 —Studies from more than a single center (or from more than a single investigator) for
613 independence of finding

614 • A showing that studies were conducted by groups with properly documented operating
615 procedures and a history of implementing such procedures effectively

616 • Prospective design to minimize bias

617 • A showing that a sufficient number of patients were enrolled in the studies to provide results
618 that are valid (we recommend this be documented by a sample size discussion in the article)

619 • A showing that there is a sufficient level of detail of information in the studies to assess:
620 —Mass dose(s) and radiation dose(s) used and regimen(s) used for dosage administration
621 —Image acquisition, including device settings, timing and interval of imaging, and views
622 obtained
623 —Image blinding method used, how images were handled and presented to the blinded
624 reader, details on sequential unblinding if used, methods used by the core laboratory before
625 images were presented to the blinded reader, and how multiple lesions were tracked
626 —Details of how the imaging agent is made so that FDA can assess the identity of an agent
627 used in multiple studies if that agent is made locally
628 --Use and description of controls to minimize bias: for example, randomization,
629 blinding, central reading versus local reading
630 --Statistical plans, prespecified analytic methods, prospectively defined study endpoints,
631 and a full accounting of the study population enrolled
632 --Study endpoints that are objective and not dependent on investigator judgment.
633 Description of imaging features used by the blinded readers to reach their decision. For
634 diagnostic tests where endpoints are interpretive, we recommend use of a well-accepted
635 truth standard, such as pathology. If study endpoints are also compared to an active
636 imaging control drug or modality, that imaging approach must be approved for the
637 indication being studied. We recommend the endpoint be clinically useful.
638 —Robust results that yield a conclusion of efficacy that is consistent with the prospective
639 protocol design and that do not require post-hoc analyses
640
641
642
643
644
645

646 **GLOSSARY**
647
648 *Note*: Subjects in trials of medical imaging agents are often classified into one of four groups
649 depending on (1) whether disease is present (often determined with a truth standard or gold
650 standard) and (2) the results of the diagnostic test of interest (positive or negative). The
651 following table identifies the variables that are used to estimate the parameters defined below.
652

Test Result:	**Disease:** Present (+)	**Disease:** Absent (-)	
Positive (+)	**TP (a)** true positive=TP	**FP (b)** false positive=FP	**m1 = a+b =TP+FP** total with positive test
Negative (-)	**FN (c)** false negative=FN	**TN (d)** true negative=TN	**m2 = c+d =FN+TN** total with negative test
	n1 = a + c = TP+FN total with disease	**n2 = b +d=FP+TN** total without disease	**N = a+b+c+d =TP+FP+FN+TN** total in study

653
654
655 **Accuracy**: (1) In common usage, *accuracy* is the quality of being true or correct. (2) As a
656 measure of diagnostic performance, *accuracy* is a measure of how faithfully the information
657 obtained using a medical imaging agent reflects reality or *truth* as measured by a truth standard
658 or *gold standard*. Accuracy is the proportion of cases, considering both positive and negative
659 test results, for which the test results are correct (i.e., concordant with the truth standard or *gold*
660 *standard*.) Accuracy = (a+d)/N = (TP+TN)/(TP+FP+FN+TN).
661
662 **Comparator:** An established test against which a proposed test is compared to evaluate the
663 effectiveness of the proposed test. A comparator usually means an agent or modality approved
664 for a similar indication. (See also definition for *reference product*.)
665
666 **Negative predictive value:** The probability that a subject does not have the disease when the
667 test result is negative. Synonyms include *predictive value negative*. Negative predictive value =
668 d/m2 = TN/(TN+FN). By application of Bayes' Rule, the negative predictive value also can be
669 defined as a function of the pretest probability of disease (p), sensitivity, and specificity:
670 negative predictive value = [(1-p) • specificity]/[(1-p) • specificity + p • (1- sensitivity)].
671
672 **Positive predictive value:** The probability that a subject has the disease when the test result is
673 positive. Synonyms include *predictive value positive*. Positive predictive value = a/m1 =
674 TP/(TP+FP). By application of Bayes' Rule, the positive predictive value also can be defined as
675 a function of pretest probability of disease (p), sensitivity, and specificity: positive predictive
676 value = [(p • sensitivity)/[p • sensitivity + (1-p) • (1- specificity)].
677
678 **Precision:** A measure of the reproducibility of a test, including reproducibility within and
679 across doses, rates of administration, routes of administration, timings of imaging after product
680 administration, instruments, instrument operators, patients, and image interpreters, and possibly

681 other variables. Precision is usually expressed in terms of variability, using such measures as
682 confidence intervals and/or standard deviations. Precise tests have relatively narrow confidence
683 intervals (or relatively small standard deviations).
684
685 **Reference product:** An FDA-approved drug product having an indication similar to that of an
686 investigational drug or biological product to which it is being compared for the purpose of
687 evaluating the effectiveness of the investigational drug or biological product.
688
689 **Sensitivity:** The probability that a test result is positive when the subject has the disease.
690 Synonyms include *true positive rate*. Sensitivity = a/n1 = TP/(TP+FN).
691
692 **Specificity:** The probability that a test result is negative when the subject does not have the
693 disease. Synonyms include *true negative rate*. Specificity = d/n2 = TN/(TN+FP).
694
695 **Truth standard (gold standard):** An independent method of measuring the same variable
696 being measured by the investigational drug or biological product that is known or believed to
697 give the *true* value of a measurement.

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 3: Design, Analysis, and Interpretation of Clinical Studies

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Guidance for Industry

Developing Medical Imaging Drug and Biological Products

Part 3: Design, Analysis, and Interpretation of Clinical Studies

Additional copies of this Guidance are available from:

Division of Drug Information HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane, Rockville, MD 20857
(Phone 301-827-4573)
Internet: http://www.fda.gov/cder/guidance/index.htm.

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
Internet: http://www.fda.gov/cber/guidelines.htm.
Mail: the Voice Information System at 800-835-4709 or 301-827-1800.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

June 2004
Clinical Medical

Table of Contents

I. INTRODUCTION **3**

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS **4**

A. Contrast Agents **4**

B. Diagnostic Radiopharmaceuticals **4**

III. GENERAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF MEDICAL IMAGING AGENTS **5**

A. Phase 1 Studies **5**

B. Phase 2 Studies **6**

C. Phase 3 Studies **7**

IV. ADDITIONAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF EFFICACY **7**

A. Selecting Subjects **7**

B. Imaging Conditions and Image Evaluations **8**

1. Imaging Conditions *8*
2. Methods and Considerations for Image Evaluation *8*
3. Steps in Image Evaluation *9*
a. Assessing objective image features 9
b. Image interpretation 10
4. Endpoints in Trials *10*
a. Image interpretations as endpoints 10
b. Objective image features as endpoints 11
c. Subjective image assessments as endpoints 11
d. Clinical outcomes as endpoints 11
5. Case Report Forms *12*
6. CRFs for Image Evaluation *12*
7. Blinded Imaging Evaluations *12*
a. Fully blinded image evaluation 13
b. Image evaluation blinded to outcome 14
c. Sequential Unblinding 15
d. Unblinded image evaluations 15
8. Independent Image Evaluations *16*
a. Consensus image evaluations 16
b. Repeated image evaluations by the same reader 16
9. Offsite and Onsite Image Evaluations *17*
10. Assessment of Interreader and Intrareader Variability *17*
11. Protocol and Nonprotocol Images *18*
a. Protocol images 18
b. Nonprotocol images 19
12. Separate or Combined Image Evaluations *19*
a. Separate image evaluations 19
b. Combined image evaluations 20

C. Truth Standards (Gold Standards) **22**

D. Comparison Groups **23**

1. Comparison to an Agent or Modality Approved for a Similar Indication *23*

a. Noninferiority studies 24

b. Agreement studies 25

2. Comparison to Placebo *26*

V. STATISTICAL ANALYSIS **26**

A. Statistical Methods **27**

B. Diagnostic Performance **28**

GLOSSARY **29**

Guidance for Industry[1]
Developing Medical Imaging Drug and Biological Products
Part 3: Design, Analysis and Interpretation of Clinical Studies

This guidance represents 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. An alternative approach may be used if such 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

This guidance is one of three guidances intended to assist developers of medical imaging drug and biological products (*medical imaging agents*) in planning and coordinating their clinical investigations and preparing and submitting investigational new drug applications (INDs), new drug applications (NDAs), biologics license applications (BLAs), abbreviated NDAs (ANDAs), and supplements to NDAs or BLAs. The three guidances are*: Part 1: Conducting Safety Assessments*; *Part 2: Clinical Indications*; and *Part 3: Design, Analysis, and Interpretation of Clinical Studies*.

Medical imaging agents generally are governed by the same regulations as other drug and biological products. However, because medical imaging agents are used solely to diagnose and monitor diseases or conditions as opposed to treat them, development programs for medical imaging agents can be tailored to reflect these particular uses. Specifically, this guidance discusses our recommendations on how to design a clinical development program for a medical imaging agent including selecting subjects and acquiring, analyzing, and interpreting medical imaging data.

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

[1] This guidance has been prepared by the Division of Medical Imaging and Radiopharmaceutical Drug Products and the Office of Therapeutics Research and Review in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration.

cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

A glossary of common terms used in diagnostic medical imaging is provided at the end of this document.

II. SCOPE — TYPES OF MEDICAL IMAGING AGENTS

This guidance discusses medical imaging agents that are administered in vivo and are used for diagnosis or monitoring with a variety of modalities, such as radiography, computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), and radionuclide imaging. The guidance is not intended to apply to the development of in vitro diagnostic or therapeutic uses of these agents.[2]

Medical imaging agents can be classified into at least two general categories:

A. Contrast Agents

As used in this guidance, a contrast agent is a medical imaging agent used to improve the visualization of tissues, organs, and physiologic processes by increasing the relative difference of imaging signal intensities in adjacent regions of the body. Types of contrast agents include (1) iodinated compounds used in radiography and CT; (2) paramagnetic metallic ions (such as ions of gadolinium, iron, and manganese) linked to a variety of molecules and microparticles (such as superparamagnetic iron oxide) used in MRI; and (3) microbubbles, microaerosomes, and related microparticles used in diagnostic ultrasonography.

B. Diagnostic Radiopharmaceuticals

As used in this guidance, a *diagnostic radiopharmaceutical* is (1) an article intended for use in the diagnosis or monitoring of a disease or a manifestation in humans and that exhibits spontaneous disintegration of unstable nuclei with the emission of nuclear particles or photons or

---

[2] The guidance is not intended to apply to the development of research drugs that do not provide direct patient benefit with respect to diagnosis, therapy, prevention, or prognosis, or other clinically useful information. These include radioactive drugs for research that are used in accordance with 21 CFR 361.1. Section 361.1 states that radioactive drugs (defined in 21 CFR 310.3(n)) are generally recognized as safe and effective when administered under specified conditions to human research subjects in the course of a project intended to obtain basic information about the metabolism of a radioactively labeled drug or about human physiology, pathophysiology, or biochemistry. However, if a radioactive drug is used for immediate therapeutic, diagnostic, or similar purpose or to determine the safety and effectiveness of the drug in humans, or if the radioactive drug has a pharmacological effect in the body, an IND is required. FDA is developing a guidance on determining when research with radioactive drugs may be conducted under § 361.1.

The Agency recognizes the potential of imaging agents as research tools for aiding the development of therapeutic drugs, and some of the principles of the guidance may be applicable to such research.. Sponsors of such imaging research agents are urged to contact the Division of Medical Imaging and Radiopharmaceutical Drug Products for advice on development of the imaging research agent.

(2) any nonradioactive reagent kit or nuclide generator that is intended to be used in the preparation of such an article.[3] As stated in the preamble to FDA's proposed rule on Regulations for In Vivo Radiopharmaceuticals Used for Diagnosis and Monitoring, the Agency interprets this definition to include articles that exhibit spontaneous disintegration leading to the reconstruction of unstable nuclei and the subsequent emission of nuclear particles or photons (63 FR 28301 at 28303; May 22, 1998).

Diagnostic radiopharmaceuticals are generally radioactive drugs or biological products that contain a radionuclide that typically is linked to a ligand or carrier.[4] These products are used in planar imaging, single photon emission computed tomography (SPECT), positron emission tomography (PET), or with other radiation detection probes.

Diagnostic radiopharmaceuticals used for imaging typically have two distinct components.

- A radionuclide that can be detected in vivo (e.g., technetium-99m, iodine-123, indium-111).

  The radionuclide typically is a radioactive atom with a relatively short physical half-life that emits radioactive decay photons having sufficient energy to penetrate the tissue mass of the patient. These photons can then be detected with imaging devices or other detectors.

- A nonradioactive component to which the radionuclide is bound that delivers the radionuclide to specific areas within the body.

  This nonradionuclidic portion of the diagnostic radiopharmaceutical often is an organic molecule such as a carbohydrate, lipid, nucleic acid, peptide, small protein, or antibody.

As technology advances, new products may emerge that do not fit into these traditional categories (e.g., agents for optical imaging, magnetic resonance spectroscopy, combined contrast and functional imaging). It is anticipated, however, that the general principles discussed here could apply to these new diagnostic products. Developers of these products are encouraged to contact the appropriate reviewing division for advice on product development.

III. GENERAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF MEDICAL IMAGING AGENTS

A. Phase 1 Studies

[3] 21 CFR 315.2 and 601.31.

[4] In this guidance, the terms *ligand* and *carrier* refer to the entire nonradionuclidic portion of the diagnostic radiopharmaceutical.

108 The general goal of phase 1 studies[5] of medical imaging agents is to obtain pharmacokinetic and
109 human safety assessments of a single mass dose and increasing mass doses of a drug or
110 biological product. We recommend that evaluation of a medical imaging agent that targets a
111 specific metabolic process or receptor include assessments of its potential effects on these
112 processes or receptors.
113
114 We recommend that, for diagnostic radiopharmaceuticals, organ and tissue distribution data over
115 time be collected to optimize subsequent imaging protocols and calculate radiation dosimetry
116 (see Part I, section IV.D). We also recommend that, as appropriate, pharmacokinetic and
117 pharmacodynamic evaluations be made of the intact diagnostic radiopharmaceutical, the carrier
118 or ligand, and other vial contents, especially when large amounts of cold components are present
119 as determined by absolute measurement or by relative concentration of labeled to unlabeled
120 carrier or ligand. This can be achieved by administering large mass doses of a medical imaging
121 agent with low specific activity, administering the contents of an entire vial of a medical imaging
122 agent (assuming that this approximates a worst-case scenario in clinical practice), or both.
123 Because of potential toxicities, this approach may not be appropriate for some drugs nor for most
124 biological products. In such cases, we recommend you contact the review division.
125
126 **B. Phase 2 Studies**
127
128 The general goals of phase 2 studies of medical imaging agents include (1) refining the agent's
129 clinically useful mass dose and radiation dose ranges or dosage regimen (e.g., bolus
130 administration or infusion) in preparation for phase 3 studies, (2) answering outstanding
131 pharmacokinetic and pharmacodynamic questions, (3) providing preliminary evidence of
132 efficacy and expanding the safety database, (4) optimizing the techniques and timing of image
133 acquisition, (5) developing methods and criteria by which images will be evaluated, and
134 (6) evaluating other critical questions about the medical imaging agent. With the
135 accomplishment of these elements, phase 3 development should proceed smoothly.
136
137 We recommend that sponsors explore the consequences of both mass dose and radiation dose (or
138 dosage regimen) adjustment on image acquisition and on the safety or effectiveness of the
139 administered product. We recommend that additional exploration include adjusting the
140 following if relevant:
141
142 • Character and amount of active and inactive ingredients
143 • Amount of radioactivity
144 • Amount of nonradioactive ligand or carrier
145 • Specific activity
146 • Radionuclide that is used
147

[5] See also the guidance *Content and Format of Investigational New Drug Applications (INDs) for Phase-1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-Derived Products*. This and all other guidances cited in this document are available at FDA's Web site at http://www.fda.gov/cder/guidance/index.htm.

148 We recommend that methods used to determine the comparability, superiority, or inferiority of
149 different mass and radiation doses or regimens be discussed with the Agency. To the extent
150 possible, the formulation that will be used for marketing should be used during phase 2 studies.
151 When a different formulation is used, we recommend that bioequivalence and/or other bridging
152 studies be used to document the relevance of data collected with the original formulation.
153
154 We recommend that phase 2 studies be designed to define the appropriate patient populations
155 and clinical settings for phase 3 studies. To gather preliminary evidence of efficacy, however,
156 both subjects with known disease (or patients with known structural or functional abnormalities)
157 and subjects known to be normal for these conditions may be included in clinical studies.
158 However, for products that are immunogenic or exhibit other toxicities, use of healthy subjects
159 may not be appropriate. We recommend that methods, endpoints, and items on the case report
160 form (CRF) that will be used in critical phase 3 studies be tested and refined.
161
162 **C. Phase 3 Studies**
163
164 The general goals of phase 3 efficacy studies for medical imaging agents include confirming the
165 principal hypotheses developed in earlier studies, demonstrating the efficacy and continued
166 safety of the medical imaging agent, and validating instructions for use and for imaging in the
167 population for which the agent is intended. We recommend that the design of phase 3 studies
168 (e.g., dosage, imaging techniques and times, patient population, and endpoints) be based on the
169 findings in phase 2 studies. We recommend that the formulation intended for marketing be used,
170 or bridging studies be performed.
171
172 When multiple efficacy studies are performed, the studies can be of different designs.[6] To
173 increase the extent to which the results can be generalized, we recommend the studies be
174 independent of one another and use different investigators, clinical centers, and readers that
175 perform the blinded image evaluations (see section IV.B).
176
177
178 **IV. ADDITIONAL CONSIDERATIONS IN THE CLINICAL EVALUATION OF**
179 **EFFICACY**
180
181 The following sections describe special considerations for the evaluation of efficacy in clinical
182 trials for medical imaging agents (see *Part 2: Clinical Indications,* section IV, for
183 recommendations on general considerations for establishing effectiveness, clinical usefulness,
184 and clinical setting).
185
186 **A. Selecting Subjects**
187
188 We recommend that subjects included in phase 3 clinical efficacy studies be representative of the
189 population in which the medical imaging agent is intended to be used. We also recommend that
190 the protocol and study reports specify the method by which patients were selected for

[6] See the guidance *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products*.

191 participation in the study (e.g., consecutive subjects enrolled, random selection) to facilitate
192 assessments of potential selection bias (e.g., using a comparator test result to pre-select subjects
193 most likely to have the desired image finding).[7]
194
195 **B. Imaging Conditions and Image Evaluations**
196
197 The following guidance may be customized to the specific medical imaging drug, biological
198 product, or imaging modality under development. (The term *images* is nonspecific and may refer
199 to an individual image or to a set of images acquired from different views, different sequences
200 and timing.)
201
202 *1. Imaging Conditions*
203
204 We recommend that the effects of changes in relevant imaging conditions (e.g., timing of
205 imaging after product administration, views, instrument settings, patient positioning) on
206 image quality and reproducibility, including any limitations imposed by changes in such
207 conditions, be evaluated in early product development. We recommend that subsequent,
208 phase 3 efficacy trials substantiate and possibly refine these conditions for use.
209 Appropriate imaging conditions, including limitations, can be described in the product
210 labeling.
211
212 *2. Methods and Considerations for Image Evaluation*
213
214 We recommend that methods and criteria for image evaluation (including criteria for
215 image interpretation) be evaluated in early product development. Subsequently, we
216 recommend that the methods and criteria that are anticipated for clinical use be employed
217 and substantiated in the phase 3 efficacy trials. For example, early clinical trials might
218 compare ways in which regions of interest on images are selected or ways in which an
219 organ will be subdivided on images for purposes of analysis. Similarly, early clinical
220 trials might evaluate which objective image features (e.g., lesion conspicuity, relative
221 count rate density) appear to be most affected by the medical imaging agent and which of
222 these are most useful in image interpretation, such as making a determination of whether
223 a mass is benign or malignant (see section IV.B.3).
224
225 We recommend that the most appropriate of these methods and criteria for image
226 evaluation be incorporated into the protocols of the phase 3 efficacy trials.

[7] To aid in the subsequent use of this information in clinical trial design, the pretest odds or pretest probabilities of disease can be used as part of the selection criteria as a method of ensuring enrollment of the population of intended use and/or as part of the patient stratification or subsetting criteria for analysis. We recommend that the range of pretest probabilities enrolled be determined by the type of clinical setting that will support the labeling (e.g., a screening setting, a case finding setting, a pivotal decision setting). We recommend that the pretest odds or probabilities be estimated for all subjects after enrollment, but before any trial results are made available. We also recommend that these odds and probabilities be derived from prespecified criteria for disease (e.g., history, physical findings, results of other diagnostic evaluations) according to prespecified algorithms. We recommend that the estimated pretest odds and probabilities of disease should be compared with the pretest odds and probabilities actually observed in the studies. (See the glossary for the definition of terms relating to pretest odds and probabilities for study analysis.)

227
228 A description of the appropriate methods and criteria for image evaluation, including
229 limitations, should be described in the product labeling.
230
231 We recommend that sponsors seek FDA comment on the designs and analysis plans for
232 the principal efficacy trials before they are finalized. In some cases, special protocol
233 assessments may be appropriate (see guidance for industry *Special Protocol Assessment*).
234 In addition, we recommend that the following elements be completed and submitted to
235 the IND before the phase 3 efficacy studies enroll subjects:
236
237 • Proposed indications for use
238 • Protocols for the phase 3 efficacy trials
239 • Investigators' brochure
240 • CRFs to be used by on-site investigators
241 • Plan for blinded image evaluations[8]
242 • CRFs to be used by the blinded readers
243 • Statistical analysis plan
244 • Plan for on-site image evaluation and intended use of such evaluation in patient
245 management, if any
246
247 We recommend that sponsors submit a single comprehensive statistical analysis plan for
248 each principal efficacy study. We recommend that this statistical analysis plan be part of
249 the study protocol, include the plan for blinded image evaluations, and be submitted to
250 the protocol before images have been collected.
251
252 *3. Steps in Image Evaluation*
253
254 The evaluation of medical images generally consists of two distinct steps: assessing
255 objective image features and interpreting findings on the image.
256
257 a. Assessing objective image features
258
259 As used in this guidance, *objective image features* are attributes on the image that
260 are either visually perceptible or that can be detected with instrumentation.
261 Examples of objective image features include signal-to-noise ratios; degree of
262 delineation; extent of opacification; and the size, number, or density of lesions.
263
264 Objective image features can be captured on scales that are continuous (e.g., the
265 diameter of a mass), ordinal (e.g., a feature can be classified as definitely
266 increased, probably increased, neither increased nor decreased, probably
267 decreased, definitely decreased), or dichotomous (e.g., a feature can be classified
268 as present or absent).
269

[8] *Blinded* image evaluations may also be referred to as *masked* or as *uninformed* image evaluations.

Medical imaging agents have their intended effects by altering objective image features. We recommend that both the nature and location of such changes on the image be documented fully during image evaluations in clinical trials intended to demonstrate efficacy. We also recommend that such documentation also include changes that are unintended or undesirable. For example, a diagnostic radiopharmaceutical intended for cardiac imaging also might localize in the liver, thereby obscuring visualization of parts of the heart.

When possible, it is often desirable to perform both a qualitative visual evaluation of images as well as a quantitative analysis of images with instrumentation. However, a quantitative image analysis with instrumentation by itself may not be sufficient to establish efficacy of the medical imaging agent, such as in cases where images are not intended (or not likely) to be evaluated quantitatively with instrumentation in clinical practice.

b. Image interpretation

As used in this guidance, an *image interpretation* is the explanation or meaning that is attributed to objective image features. We recommend that interpretations of image features be supported by objective, quantitative, and/or qualitative information derived from the images. For example, the interpretation that cardiac tissue seen on an image is infarcted, ischemic, or normal might be supported by objective image features such as the extent and distribution of localization of the medical imaging agent in the heart (e.g., increased, normal, decreased, or absent), the time course of such localization, and how these features are affected by exercise or pharmacologic stress.

4. Endpoints in Trials

Medical imaging agents could be developed for structural delineation; functional, physiological, or biochemical assessment; disease or pathology detection or assessment; diagnostic or therapeutic patient management; or multiple or other indications. The primary endpoints (response variables) relate to the indication's clinical usefulness (see Part 2: Clinical Indications, section IV.B).

a. Image interpretations as endpoints

Image interpretations that are clinically useful can be incorporated into the primary endpoint in phase 3 clinical trials. For example, the primary analysis endpoints of a trial for a medical imaging agent intended for the indication *disease or pathology detection or assessment* might be the proportions of subjects with and without the disease who are properly classified against an appropriate truth standard. In this example, the interpretation that a pulmonary lesion seen on an image is benign or malignant has direct clinical meaning and can be incorporated into the primary endpoint.

316 b. Objective image features as endpoints
317
318 When the clinical usefulness of particular objective image features is obvious and
319 apparent, the objective imaging features can be incorporated into the primary
320 endpoint. For example, in a study of a medical imaging agent intended for brain
321 imaging, the ability to delineate anatomy that indicates the presence or absence of
322 cranial masses on images has direct clinical usefulness. The primary endpoint
323 (e.g., cranial mass detection) serves as the primary basis for the indication for the
324 product (e.g., the medical imaging agent is indicated for detecting cranial masses
325 in patients in a particular defined clinical setting).
326
327 However, in some cases the clinical usefulness of particular objective image
328 features may not be readily apparent without additional interpretation. In these
329 cases, we recommend that the objective image features serve as secondary
330 imaging endpoints. For example, the finding that a medical imaging agent alters
331 the conspicuity of masses differentially could lead to the interpretation that
332 specific masses are benign or malignant; acute or chronic; inflammatory,
333 neoplastic, or hemorrhagic; or lead to some other clinically useful interpretations.
334 The interpretations can be incorporated into the primary endpoint and can serve as
335 the primary basis for the indication for the product. However, the objective image
336 feature of lesion conspicuity might be designated more appropriately as a
337 secondary imaging endpoint.
338
339 c. Subjective image assessments as endpoints
340
341 As used in this guidance, *subjective image assessments* are perceptions or
342 inferences made by the reader. Such assessments are tangible and cannot be
343 measured objectively. For example, a conclusion that use of a medical imaging
344 agent alters *diagnostic confidence* is a subjective assessment as is the conclusion
345 that a medical imaging agent provides *more diagnostic information*.
346
347 We recommend that subjective image assessments be linked to objective image
348 features so that the objective basis for such assessments can be understood.
349 Subjective image assessments can be difficult to validate and replicate. They may
350 introduce bias as well. Therefore, subjective image assessments should not be
351 used as primary imaging endpoints.
352
353 d. Clinical outcomes as endpoints
354
355 Clinical outcomes, such as measurement of symptoms, functioning, or survival,
356 are among the most direct ways to measure clinical usefulness. Clinical outcomes
357 can serve as primary endpoints in trials of medical imaging agents. For example,
358 the primary endpoint of a trial of a medical imaging agent intended for the
359 indication *therapeutic patient management* in patients with colon cancer might be
360 a response variable that measures changes in symptoms, functioning, or survival.
361

362 *5. Case Report Forms*
363
364 We recommend that case report forms (CRFs) in trials of medical imaging agents
365 prospectively define the types of observations and evaluations for investigators to record.
366 In addition to data that are usually recorded in CRFs (e.g., inclusion/exclusion criteria,
367 safety findings, efficacy findings), we recommend that the onsite investigator's CRF for a
368 medical imaging agent capture the following information:
369
370 • The technical performance of the diagnostic radiopharmaceutical used in the
371 study, if any (e.g., specific activity, percent bound, percent free, percent
372 active, percent inactive)
373
374 • The technical characteristics and technical performance of the imaging
375 equipment (e.g., background flood, quality control analysis of the imaging
376 device, pulse height analyzer)
377
378 • Methods of image acquisition, output processing, display, reconstruction, and
379 archiving of the imaging study
380
381 The collection and availability of the data on the CRF may be important for labeling how
382 the imaging agent is intended to be administered and the appropriate device settings for
383 optimal imaging.
384
385 *6. CRFs for Image Evaluation*
386
387 We recommend that imaging CRFs be designed to capture imaging endpoints, including
388 objective features of the images as well as the location and interpretation of any findings.
389 We recommend that interpretations of image features be supported by objective
390 quantitative or qualitative information derived from the images. We recommend that
391 image interpretations be recorded as distinct items from the assessments of the objective
392 image features. We also recommend that items on the CRFs for image evaluation be
393 carefully constructed to gather information without introducing a bias that suggests the
394 answer that is being sought. We recommend that the proposed labeled indication be
395 clearly derived from specific items in the CRF and from endpoints and hypotheses that
396 have been prospectively stated in the protocol.
397
398 *7. Blinded Imaging Evaluations*
399
400 We recommend that image evaluations be designed to demonstrate that the specific
401 effects of the medical imaging agent, as manifested in the images, provide such
402 information reproducibly and apart from other possible confounding influences or biases.
403 We recommend that blinded image evaluations by multiple independent readers be
404 performed in the phase 3 efficacy studies.
405
406 We recommend that either a *fully blinded image evaluation* or an *image evaluation*
407 *blinded to outcome* by independent readers serve as the principal image evaluation for

demonstration of efficacy.[9] Alternatively, both types of image evaluations can be used; if so, the evaluations can be performed through sequential unblinding. Both primary and secondary imaging endpoints should be evaluated in this manner. We recommend that the nature and type of information available to the readers be discussed with FDA before the trials are initiated.

In addition to the items outlined in the sections below, we recommend that plans for blinded image evaluations include the following elements:

- We recommend that the protocol clearly specify the elements to which readers are blinded.

- We recommend that meanings of all endpoints be clearly understood for consistency. We recommend that terms to be used in image evaluation and classification be defined explicitly in the image evaluation plan, including such terms as *technically inadequate*, *uninterpretable*, *indeterminate*, or *intermediate*. Blinded readers can be trained in scoring procedures using sample images from phase 1 and phase 2 studies.

- We recommend that images be masked for all patient identifiers.

- We recommend that blinded readers evaluate images in a random sequence. *Randomization* of images refers to merging the images obtained in the study (to the fullest degree that is practical) and then presenting images in this merged set to the readers in a random sequence.

  For example, when images of several diagnostic radiopharmaceuticals read by the same criteria are being compared to establish relative efficacy (e.g., a comparison of a test drug or biological product to an established drug or biological product), we recommend the readers evaluate individual images from the merged set of images in a random sequence.

  a. Fully blinded image evaluation

  During a *fully blinded image evaluation*, we recommend that readers not have any knowledge of the following types of information:

  - Results of evaluation with the truth standard, of the final diagnosis, or of patient outcome

  - Any patient-specific information (e.g., history, physical exam, laboratory results, results of other imaging studies)

[9] See section IV.B.8 for a definition of *independent readers*.

450 We recommend that general inclusion and exclusion criteria for patient
451 enrollment, other details of the protocol, or anatomic orientation to the images not
452 be provided to the readers.
453
454 During a *fully blinded image evaluation* in studies where images obtained by
455 different treatments are being evaluated, we recommend that readers not have
456 knowledge of treatment identity, to the greatest extent to which that is possible.[10]
457 For example, in a comparative study of two or more medical imaging agents (or
458 of two or more doses or regimens of a particular medical imaging agent), we
459 suggest the blinded readers not know which agent (or which dose or regimen) was
460 used to obtain a given image.
461
462 For contrast agents, we suggest this also can include lack of knowledge about
463 which images were obtained before product administration and which were
464 obtained after product administration, although sometimes this is apparent upon
465 viewing the images.
466
467 In cases where the instructions for image evaluation differ according to treatment
468 (e.g., as might be the case when images are obtained using different imaging
469 modalities), blinding the readers to treatment identity may be infeasible.
470
471 b. Image evaluation blinded to outcome
472
473 As in a *fully blinded image evaluation*, we recommend that readers performing an
474 *image evaluation blinded to outcome* not have any knowledge of the results of
475 evaluation with the truth standard, of the final diagnosis, or of patient outcome.
476
477 However, in an *image evaluation blinded to outcome,* the readers might have
478 knowledge of particular elements of patient-specific information (e.g., history,
479 physical exam, laboratory results, or results of other imaging studies). In some
480 cases, the readers also might be aware of general inclusion and exclusion criteria
481 for patient enrollment, other details of the protocol, or anatomic orientation to the
482 images. We recommend that the particular elements about which the reader will
483 have information be standardized for all patients and defined prospectively in the
484 clinical trial protocol, statistical plan, and the blinded image evaluation plan.
485
486 In studies where images obtained by different treatments are being evaluated
487 (including *no treatment*, such as in unenhanced image evaluation of a contrast
488 agent), we recommend that the readers not have knowledge of treatment identity,
489 to the greatest extent to which that is possible (see section IV.B.7.a).
490

[10] This is the common meaning of *blinding* in therapeutic clinical trials. See the ICH guidelines *E8 General Considerations for Clinical Trials* and *E9 Statistical Principles for Clinical Trials*.

491 c. Sequential Unblinding
492
493 As used in this guidance, *sequential unblinding* is an assessment where readers
494 typically evaluate images with progressively more information (e.g., clinical
495 information) on each read. Sequential unblinding might be used to provide
496 incremental information under a variety of conditions that may occur in routine
497 clinical practice (e.g., when no clinical information is available, when limited
498 clinical information is available, and when a substantial amount of information is
499 available). This can be used to determine when or how the test agent should be
500 used in a diagnostic algorithm. We recommend that a typical s*equential*
501 *unblinding* image evaluation be a three-step process.
502
503 • We recommend that a fully blinded image evaluation be performed. We
504 recommend that this evaluation be recorded and locked in a dataset by
505 methods that can be validated. In a *locked* dataset, we recommend that it not
506 be possible to alter the evaluation later when additional information is
507 available, or if input is received from the clinical investigators, other readers,
508 or the sponsor.
509 • We recommend that an image evaluation blinded to outcome be performed.
510 We recommend this evaluation be recorded and locked in the dataset.
511 • To determine diagnostic performance of the imaging agent, we recommend
512 that the result of the above two blinded evaluations be compared to the results
513 of evaluation with the truth standard (or of the final diagnosis, or of patient
514 outcome).
515
516 Such sequential unblinding can be expanded to include other types of image
517 evaluations where additional clinical information is provided to the readers. If
518 sequential unblinding is used, we recommend that the protocol specify the
519 hypothesis that is to be evaluated at each step. Also, we recommend that the
520 protocol specify which image evaluation will be the primary one for determining
521 efficacy.[11]
522
523 d. Unblinded image evaluations
524
525 In an *unblinded image evaluation,* readers are aware of the results of patient
526 evaluation with the truth standard, of the final diagnosis, or of patient outcome.
527 Unblinded readers also typically are aware of patient-specific information
528 (e.g., history, physical exam, laboratory results, results of other imaging studies),
529 of treatment identity where images obtained by different treatments (including no
530 treatment) are being evaluated, of inclusion and exclusion criteria for patient

[11] The labeling should reflect the image methods (blinded, sequentially unblinded, or unblinded, as appropriate) that provided substantial evidence that the Agency used to reach an approval decision and to develop appropriate labeling recommendations for use.

enrollment, other details of the protocol, and of anatomic orientation to the images.

Unblinded image evaluations can be used to show consistency with the results of fully blinded image evaluations or image evaluations blinded to outcome. We recommend that these blinded and unblinded image evaluations use the same endpoints so that the results can be compared. However, we recommend that unblinded image evaluations not be used as the principal image evaluation for demonstration of efficacy. The unblinded readers may have access to additional information that may alter the readers' diagnostic assessments and may confound or bias the image evaluation by these readers.

8. Independent Image Evaluations

Two events are independent if knowing the outcome of one event says nothing about the outcome of the other. Therefore, as used in this guidance, *independent readers* are readers that are completely unaware of findings of other readers (including findings of other blinded readers and onsite investigators) and are readers who are not otherwise influenced by the findings of other readers. To ensure that blinded reader's evaluations remain independent, we recommend that each blinded reader's evaluation be locked in the dataset shortly after it is obtained and before additional types of image evaluations are performed (see section IV.B.7.c).

a. Consensus image evaluations

As used in this guidance, c*onsensus image evaluations* (*consensus reads*) are image evaluations during which readers convene to evaluate images together. Consensus image evaluations can be performed after the individual readings are completed and locked. However, readers are not considered independent during consensus reads and therefore we recommend that such reads not serve as the primary image evaluation used to demonstrate the efficacy of medical imaging agents. Although a consensus read is performed by several readers, it is actually a single image-evaluation and is unlikely to fulfill our interest in image evaluations by multiple blinded readers. As with the individual blinded evaluations, we recommend that the consensus reads be locked once obtained and before additional types of blinded readings are performed.

b. Repeated image evaluations by the same reader

In studies where readers evaluate the same image multiple times (e.g., as in sequential unblinding, or in readings designed to assess *intra*reader variability), we recommend that the readings be performed independently of one another to the fullest extent practical. The goal is to minimize *recall bias*. We further recommend that readers be unaware, to the fullest extent practical, of their own previous image findings and not be otherwise influenced by those previous findings.

577
578 We recommend that different pages in the CRF be used for the two image
579 evaluations and that each image evaluation be performed with sufficient time
580 between readings to decrease recall and without reference to prior results.
581
582 *9. Offsite and Onsite Image Evaluations*
583
584 As used in this guidance, *offsite image evaluations* are image evaluations performed at
585 sites that have not otherwise been involved in the conduct of the study and by readers
586 who have not had contact with patients, investigators, or other individuals involved in the
587 study. We recommend that Phase 3 trials include offsite image evaluations that are
588 performed at a limited number of sites (or preferably at a centralized site). In such offsite
589 evaluations, it is usually easier to control factors that can compromise the integrity of the
590 blinded image evaluations and to ensure that the blinded readers perform their image
591 evaluations independently of other image evaluations.
592
593 As used in this guidance, o*nsite image evaluations* are image evaluations performed by
594 investigators involved in the conduct of the protocol or in the care of the patient. The
595 term also can refer to blinded image evaluations performed at sites involved with the
596 conduct of the study. Onsite investigators may have additional information about the
597 patients that was not predefined in the clinical trial protocol. Such additional information
598 may alter the investigators' diagnostic assessments and may confound or bias the image
599 evaluation by the investigators. Therefore, we recommend that onsite image evaluations
600 usually not be used as the principal image evaluation for demonstration of efficacy, but
601 be regarded as supportive of the blinded image evaluations.
602
603 However, we suggest onsite investigators who are blinded to *truth* (e.g., blinded to any
604 test result that makes up the truth standard, to the final diagnosis, and to patient final
605 outcome as in an image evaluation blinded to outcome see (section IV.B.7.b)) can be
606 used for principal image evaluation. In such instances, we recommend that all clinical
607 information available to the investigator at the time of the image evaluation be clearly
608 specified and fully documented. We also recommend that a critical assessment of how
609 such information might have influenced the readings be performed. In addition, we
610 recommend that an independent blinded evaluation that is supportive of the finding of
611 efficacy be performed.
612
613 *10. Assessment of Interreader and Intrareader Variability*
614
615 We recommend that at least two blinded readers (and preferably three or more) evaluate
616 images for each study that is intended to demonstrate efficacy. (The truth standard,
617 however, may be read by a single blinded reader.) The use of multiple readers allows for
618 an evaluation of the reproducibility of the readings (i.e., interreader variability) and
619 provides a better basis for subsequent generalization of any findings. Ideally, we
620 recommend that each reader view all of the images intended to demonstrate efficacy,
621 both for the investigational imaging agent and the truth standard, so that interreader
622 agreement can be measured. In large studies, where it may be impractical to have every

623 image read by each reader, a properly chosen subset of images can be selected for such
624 duplicate image evaluations. We recommend that consistency among readers be
625 measured quantitatively (e.g., with the kappa statistic).
626
627 We recommend that *intra*reader variability be assessed during the development of
628 medical imaging agents. This can be accomplished by having individual blinded readers
629 perform repeated image evaluations on some or all images (see section IV.B.8.b).
630
631 *11. Protocol and Nonprotocol Images*
632
633 Images obtained in a clinical trial of a medical imaging agent can generally be considered
634 either protocol or nonprotocol images.
635
636 a. Protocol images
637
638 As used in this guidance, *protocol images* are images obtained under protocol-
639 specified conditions and at protocol-specified time points with the goal of
640 demonstrating or supporting efficacy. We recommend that efficacy evaluations
641 be based on the evaluations of such protocol images. We also recommend that all
642 protocol images (e.g., not just those images determined to be evaluable) be
643 evaluated by the blinded readers, including images of test patients, control
644 patients, and normal subjects. In addition, we recommend that evaluation of the
645 protocol images be completed before other images, such as nonprotocol images,
646 are reviewed by the readers (see section IV.B.11.b).
647
648 In some cases where large numbers of images are obtained or where image tapes
649 are obtained (e.g., cardiac echocardiography), sponsors have used image selection
650 procedures. This is discouraged because the selection of images can introduce the
651 bias of the selector.
652
653 We recommend that sponsors specify prospectively in protocols of efficacy
654 studies how missing images (and images that are technically inadequate,
655 uninterpretable or show results that are indeterminate or intermediate) will be
656 handled in the data analysis. Sponsors are encouraged to incorporate analyses in
657 the statistical analysis plan that incorporate the principle of *intention-to-treat*, but
658 that are adapted to a diagnostic setting (e.g., *intention-to-diagnose* considers all
659 subjects enrolled in a diagnostic study regardless of whether they were imaged
660 with the test drug and regardless of the image quality).[12] Images (including truth
661 standard images) may be missing from analysis for many reasons, including
662 patient withdrawal from the study, technical problems with imaging, protocol

[12] The *intention-to-treat principle* is defined as the principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject (i.e., the planned treatment regimen) rather than the actual treatment given. As a consequence, we recommend that subjects allocated to a treatment group be followed up, assessed, and analyzed as members of that group irrespective of their compliance with the planned course of treatment (see *E9 Statistical Principles for Clinical Trials*, p. 28).

violations, and image selection procedures. We suggest that appropriate methods be prospectively developed to deal with missing values in the primary response variable analysis.[13]

b. Nonprotocol images

As used in this guidance, *nonprotocol image* refers to an image that is not a protocol image, as defined above (see section IV.B.11.a). These are sometimes obtained for exploratory purposes and are excluded from the locked phase 3 datasets.

12. Separate or Combined Image Evaluations

Performance of a separate image evaluation does not preclude performance of a combined image evaluation, and vice versa. If multiple image evaluations are performed, however, we recommend that the protocol specify which image evaluation will serve as the primary evaluation and which image evaluations are secondary.

a. Separate image evaluations

As used in this guidance, a *separate* image evaluation has a reader evaluate test images obtained from a patient independently of other test images obtained from that patient, to the fullest degree practical.[14] A reader evaluates each test image for a patient on its own merits without reference to, or recall of, any other test images obtained from that patient, to the fullest degree practical.

A separate image evaluation often can be performed by combining test images obtained under different conditions (or at different times) into an intermixed set. Images in this intermixed set can then be evaluated individually in random order so that multiple images are not viewed simultaneously, and so that images are not evaluated sequentially within patients. Alternatively, test images obtained under one condition (or at a particular time) can be evaluated individually in a random order, followed by an evaluation in random order of the individual test images obtained under different conditions (or at different times).

As described in the first example below, we recommend that an appropriately designed separate image evaluation be performed when a goal of a study is to make comparative inferences about product performance (e.g., to compare the diagnostic performance of one medical imaging agent with another). As described in the second example, an appropriately designed separate image evaluation also can be used to demonstrate that a contrast agent contributes additional information to images obtained with the device alone.

[13] See *E9 Statistical Principles for Clinical Trials*, p. 31.

[14] In the special case where only two test images are being evaluated, a *separate* image evaluation may also be referred to as an *unpaired* image evaluation.

705
706 Example 1: Comparative inferences of product performance
707
708 In a comparative study designed to show that the diagnostic performance of a new
709 medical imaging agent is superior to that of an approved agent and that the new
710 agent can replace the approved agent (see section IV.D.1), we recommend that an
711 appropriate separate image evaluation of test images be performed as the principal
712 image analysis. The *test images* in this case are the images obtained with the new
713 and the approved medical imaging agents. The two agents are not intended to be
714 used together in actual clinical practice, and we therefore recommend that the
715 goal of such an *unpaired* image evaluation be to show that the information
716 obtained with the new agent is clinically and statistically superior to the
717 information obtained with the approved agent. For any given patient, we
718 recommend that images obtained with the new agent be evaluated independently
719 of the evaluation of the images obtained with the approved agent, to the fullest
720 degree practical.
721
722 If desired, a side-by-side (*paired)* comparison of images obtained with the new
723 agent and the approved agent can be performed as a secondary image analysis.
724 However, such a side-by-side comparison may yield estimates of diagnostic
725 performance that are biased. The blinded reader may tend to *overread* the
726 presence of masses on the image obtained with the new agent in such a paired
727 comparison. Similarly, the blinded reader may tend to *underread* the image
728 obtained with the new agent in a paired evaluation where a mass is not seen
729 clearly on the image obtained with the approved agent.
730
731 In general, these procedures for image evaluation also are applicable to studies
732 designed to show noninferiority. We recommend that sponsors seek Agency
733 comment on proposed study designs and analytical plans before enrolling patients
734 in such studies (see also section IV.D.1 for additional discussion).
735
736 Example 2: Contribution of additional information by a contrast agent
737
738 In a study intended to demonstrate that a contrast agent contributes additional
739 information to images obtained with the device alone, it is often highly desirable
740 to perform an appropriate separate image evaluation of test images as the
741 principal image analysis (see the next section for an alternative approach). The
742 *test images*, in this case, include both the images obtained before administration
743 of contrast (the *unenhanced* images) and those obtained after administration of
744 contrast (the *enhanced* images). We recommend that the goal of such an unpaired
745 image evaluation be to show that the information obtained from the enhanced
746 image is clinically and statistically superior to the information obtained from the
747 unenhanced image.
748
749 b. Combined image evaluations
750

751 As used in this guidance, a *combined* image evaluation has a reader
752 simultaneously evaluate two or more test images that were obtained under
753 different conditions or at different times with respect to agent administration.[15] A
754 combined image evaluation may resemble the conditions under which the product
755 will be used clinically. For example, in some clinical situations both unenhanced
756 and enhanced imaging studies are typically performed in patients.[16] If so, such
757 images often are evaluated concurrently in a comparative fashion.[17] However, as
758 noted above, such combined image evaluations may increase the likelihood that
759 bias will be introduced into the image evaluations (e.g., by systematic overreading
760 or underreading particular findings on images).
761
762 A combined image evaluation can be performed by creating a set of combined
763 images for each patient. These sets can then be presented to the blinded readers
764 in random sequence.
765
766 When this type of reading is performed, however, we recommend that an
767 additional independent *separate* image evaluation be completed on at least one of
768 the members of the combination. We recommend that the member chosen be the
769 member that usually is obtained under the current standard of practice (e.g., the
770 unenhanced image). In this way, differences in the evaluations of the combined
771 reading with those of the separate reading can be assessed. When the goal is to
772 show that the medical imaging agent adds information to images, we suggest that
773 these differences demonstrate that the information from the combined images is
774 clinically and statistically superior to information obtained from the separate
775 image alone. The results of the combined and separate image evaluations can be
776 analyzed statistically using paired comparisons.
777
778 For example, when a two-dimensional ultrasound study of blood vessels is
779 performed with a microbubble contrast agent, a combined image evaluation could
780 be performed by evaluating for each patient the unenhanced and enhanced images
781 side-by-side (or in close temporal proximity). A separate independent evaluation
782 of the unenhanced image of the blood vessel (i.e., images obtained with the
783 device alone) for each patient could also be performed. Assessing the differences
784 for each patient between the results of the combined reading with those of the
785 separate readings could allow the effects of the microbubble on the images to be
786 determined.

---

[15] In the special case where only two test images are being evaluated, a *combined* image evaluation can also be referred to as a *paired* image evaluation.

[16] Also, combined images may refer to results from the test drug and modality plus images from a different modality.

[17] Under sections 505 and 502 of the Act, if images are evaluated only in a combined fashion, the approved labeling of the medical imaging agent likely will have to specify that combined evaluations should be performed in clinical practice. If such labeling restrictions are not desired, we recommend that additional separate image evaluations be performed.

As noted above, we recommend that combined and separate image evaluations be performed independently of one another to decrease recall bias (see section IV.B.8.b). We recommend that different pages in the CRF be used for the combined and separate evaluations and that the combined and separate image evaluations be performed at different times without reference to prior results.

We recommend that when differences between the combined and separate images are to be assessed, the combined CRF and separate CRF contain items or questions that are identical so that differences can be calculated and biases can be reduced by avoiding questions asking for comparative judgment.

C. Truth Standards (Gold Standards)

A truth standard provides an independent way of evaluating the same variable being assessed by the investigational medical imaging agent. A truth standard is known or believed to give the true state of a patient or true value of a measurement. Truth standards are used to demonstrate that the results obtained with the medical imaging agent are valid and reliable and to define summary test statistics (e.g., sensitivity, specificity, positive and negative predictive value). We recommend that the following general principles be incorporated prospectively into the design, conduct, and analysis of the phase 3 efficacy trials for medical imaging agents:

1. We recommend that the test results obtained with the medical imaging agent be evaluated without knowledge of the results obtained with the truth standard and without knowledge of outcome (see section IV.B.7).

2. We recommend that the true state of the subjects (e.g., diseased or nondiseased) be determined with a truth standard without knowledge of the test results obtained with the medical imaging agent.

3. We recommend that truth standards not include as a component any test results obtained with the test medical imaging agent (i.e., to avoid *incorporation bias)*. This is because the features of the test image obtained with the test agent (e.g., the *enhanced image*) are likely to be correlated to the features of the image obtained with the device alone (e.g., the *unenhanced image*). For example, in the case of a CT contrast agent intended to visualize abdominal masses, unenhanced abdominal CT images should not be included in the truth standard. However, components of the truth standard might include results from other imaging modalities (e.g., MRI, ultrasonography).

4. We recommend that evaluation with the truth standard be planned for all enrolled subjects, and the decision to evaluate a subject with the truth standard not be affected by the test results with the medical imaging agent under study. For example, if patients with positive results with the test agent are evaluated preferentially with the truth standard (as compared to patients with negative test results), the results of the study may be affected by *partial verification bias*. Similarly, if patients with positive results with the test agent are evaluated preferentially with the truth standard and those with negative test results are

833 evaluated preferentially with a less rigorous standard, the results of the study may be
834 affected by *differential verification bias.*[18]
835
836 We encourage sponsors to seek FDA comment when it is anticipated that a meaningful
837 proportion of enrolled subjects might not be evaluated with the truth standard or might be
838 evaluated with a less rigorous standard. In such situations, it may be appropriate to
839 evaluate clinical outcomes for the enrolled subjects (see section IV.D.4).
840
841 From a practical perspective, diagnostic standards are derived from procedures that are
842 considered more definitive in approximating the truth than the test agent. For
843 example, histopathology or long-term clinical outcomes may be acceptable diagnostic standards
844 for determining whether a mass is malignant. Diagnostic standards may not be error free, but for
845 purposes of the clinical trial, they generally are regarded as definitive. However,
846 misclassification of disease by the truth standard can lead to positive or negative biases in
847 diagnostic performance measures (*misclassification bias*). Thus, we recommend that the choice
848 of the truth standard be discussed with the Agency during design of the clinical trials to ensure
849 that it is appropriate.
850
851 After the truth standard has been selected, we recommend that the hypothesis for the summary
852 test statistic in reference to the truth standard be determined and prospectively incorporated into
853 the study protocol. We recommend that the hypothesis and expected summary statistics reflect
854 the intended clinical setting for use of the imaging agent (e.g., screening test, sequential
855 evaluation, alternative to or replacement of another imaging study (see section V)).
856
857 **D. Comparison Groups**
858
859 Before selecting comparison groups, discussions with the Agency are recommended. General
860 principles relating to the choice of control groups in clinical trials are set forth in the ICH
861 guideline *E10 Choice of Control Group and Related Issues in Clinical Trials* (ICH *E10*), and
862 these principles are applicable to diagnostic trials.
863
864 *1. Comparison to an Agent or Modality Approved for a Similar Indication*
865
866 If the test agent is being developed as an advance over an approved drug, biological
867 product, or other diagnostic modality, we recommend that a direct, concurrent
868 comparison to the approved comparator(s) be performed. We recommend that the
869 comparison include an evaluation of both the safety and the efficacy data for the
870 comparator(s) and the test agent. Because of disease variability, typically such
871 comparisons are performed in the same patient. We recommend that the image
872 evaluation for the test product or modality be done without knowledge of the imaging
873 results obtained from the approved products or modalities (see section IV.B.7).
874

[18] Partial verification bias and differential verification bias are forms of *diagnostic work-up bias*.

875 We recommend that information from both the test and comparator images (i.e., using the
876 new and old methods) be compared not only to one another but also to an independent
877 truth standard. This will facilitate an assessment of possible differences between the
878 medical imaging agent and the comparator and will enable comparative assessments of
879 diagnostic performance. Such assessments could be obtained, for example, by comparing
880 estimates of sensitivity, specificity, positive and negative predictive values, likelihood
881 ratios, related measures, or receiver operating characteristic (ROC) curves for the
882 different diagnostic agents. Note that two medical imaging agents could have similar
883 values for sensitivity and specificity in the same set of patients, yet have poor agreement
884 rates with each other. Similarly, two medical imaging agents could have good agreement
885 rates, yet both have poor sensitivity and specificity values. In ROC analysis, overall
886 areas under the curves obtained with different agents may be comparable, but areas under
887 partial spans of the curves may be dissimilar. Likewise, one diagnostic agent may have
888 superior diagnostic performance characteristics over another at one point on the ROC
889 curve, but may have inferior diagnostic performance characteristics at a different point
890 (see section V.B).
891
892 When a medical imaging drug or biological product is being developed for an indication
893 for which other drugs, biological products, or diagnostic modalities have already been
894 approved, a direct, concurrent comparison to the approved drug, biological product, or
895 diagnostic modality is encouraged. However, prior approval of a medical imaging agent
896 for use in a particular indication does not necessarily mean that the results of a test with
897 that agent alone can be used as a truth standard. For example, if a medical imaging agent
898 has been approved on the basis of sufficient concordance of findings with truth as
899 determined by histopathology, we recommend that assessment of the proposed medical
900 imaging agent also include determination of truth by histopathology. In this case, the
901 direct and concurrent comparison of the proposed medical imaging agent to the approved
902 agent with histopathology serving as the truth standard best measures the performance
903 difference between the two agents.
904
905 In studies that compare the effects of a test agent with another drug, biological product,
906 or imaging modality, we recommend that any images obtained using a nontest agent that
907 are taken before enrollment be used only as enrollment criteria. We recommend that
908 these images not be part of the database used to determine test agent performance. Such
909 baseline enrollment images have inherent selection bias because they are unblinded and
910 based on referral and management preferences. We recommend that test agent
911 administration be within a time frame when the disease process is expected not to have
912 changed significantly. This provides for a fair, balanced comparison between the test and
913 the comparator agent.
914
915 a. Noninferiority studies
916
917 Trials can be designed to show that a new test agent is not inferior to a reference
918 product. In general, the requirements for such studies are more stringent that the
919 requirements for studies designed to show superiority. Imaging studies, in
920 particular, can lack assay sensitivity for several reasons, including inappropriate

study population, lack of objective imaging endpoints, and inaccuracy in the truth standard. Moreover, assay sensitivity is difficult to validate because imaging studies often lack historical evidence of sensitivity to drug effects, and it is not always clear that the conduct of the imaging procedures and the subsequent image evaluations did not undermine the trial's ability to distinguish effective treatments from less effective ones. ICH *E10* provides further guidance on these matters.

We recommend that noninferiority studies be based on a concurrent comparison of the test agent and a reference product and that such studies use objectively defined endpoints validated by an acceptable truth standard. Such designs allow comparative assessment of the diagnostic (or functional) performance of the new and reference tests. For example, if the study endpoint is the presence or absence of disease, the sensitivities and specificities of the test product and the reference product can each be compared. The statistical hypotheses may be superiority, noninferiority, or both. If the test agent is to be used primarily to rule out disease, high negative predictive value and thus high sensitivity might be more important than specificity. The objective then would be to show that the new agent, when compared to the reference test, is superior with regard to sensitivity but not inferior with regard to specificity.

When the study design includes a truth standard but no comparison to a reference product, the performance levels of the new test agent can only be compared to some fixed threshold (e.g., prespecified levels of sensitivity and specificity). The statistical objective should then be to show superiority to the threshold values. Such values should be based on substantial clinical evidence supporting the assertion that exceeding the thresholds clearly demonstrates product efficacy.

To obtain a noninferiority claim against a reference product, a sponsor should show that its test agent has been shown to have similar performance characteristics as the reference product and can be used as an alternative modality in a precisely defined clinical setting. In other situations, the noninferiority comparison might only serve as a demonstration of efficacy of the test product. Generally, non-inferiority trials are designed to show that new and comparator test performance differ at most by a clinically acceptable margin that has been agreed to by the Agency. We recommend that noninferiority trials be carefully planned and that discussions with the Agency begin early in the development program.

b. Agreement studies

Similarity between a new test agent and a reference product can also be shown by demonstrating that both agents consistently give identical results. In this case, the use of a truth standard is not possible, and the objective is to show agreement between test and comparator outcomes even though the validity (accuracy) of the outcomes cannot be verified. High agreement between a new test product and a

reference product can support a claim that the new test is an acceptable alternative to the reference product.

In agreement studies, assay sensitivity is critical. In particular, outcomes should be objectively defined and the two agents should be compared in subjects who represent an appropriate spectrum of disease conditions. For example, showing that two diagnostic tests give the same positive diagnosis for a large percentage of the trial subjects might not be sufficient. We recommend that the sponsor also demonstrate that the test agent and the reference product respond similarly when a negative diagnosis prevails and that the probability of discordant outcomes is negligible. When outcomes are multivalued as opposed to dichotomous, agreement should be shown across the entire range of test values.

An agreement hypothesis should not imply that the agreement between test and comparator outcomes exceeds agreement among comparator outcomes. Thus, an understanding of intra-test and intra-reader variability should be taken into account. For example, consider a new pharmacological stress agent used with myocardial perfusion imaging to assess perfusion defects. One possible design would be to apply the comparator procedure to all subjects for a first evaluation and, for a second evaluation, randomize subjects to receive either the comparator procedure or the new test agent. This would allow the inter-test agreement to be directly compared with the intra-test agreement of the comparator using a noninferiority hypothesis.

Because agreement studies do not provide direct evidence of new test validity, they are difficult to design and execute effectively. Therefore, we recommend that sponsors pursue agreement studies in limited circumstances and consider alternative designs that employ an acceptable truth standard.

2. Comparison to Placebo

Whether the use of a placebo is appropriate in the evaluation of a medical imaging agent depends on the specific imaging agent, proposed indication, and imaging modality. In some cases, the use of placebos can help reduce potential bias in the conduct of the study and can facilitate unambiguous interpretation of efficacy or safety data. However, in some diagnostic studies (such as ultrasonography), products that are considered to be placebos (e.g., water, saline, or vehicle) can have some diagnostic effects. We recommend that these be used as controls to demonstrate that the medical imaging agent has an effect above and beyond that of its vehicle.

V. STATISTICAL ANALYSIS

We recommend that statistical methods and the methods by which diagnostic performance will be assessed be incorporated prospectively into the statistical analysis plan for each study (see section IV.B.2). In addition, we recommend that each study protocol clearly state the hypotheses

1012 to be tested, present sample size assumptions and calculations, and describe the planned
1013 statistical methods and other data analysis considerations. The ICH guideline *E9 Statistical*
1014 *Principles for Clinical Trials* provides guidance on these matters.
1015
1016 **A. Statistical Methods**
1017
1018 One part of imaging evaluation is the determination of how well the test measures what it is
1019 intended to measure (validity). The overall diagnostic performance of the product can be
1020 measured by factors such as sensitivity, specificity, positive and negative predictive values, and
1021 likelihood ratios. Outcome validity can be demonstrated by a showing that use of the test
1022 enhances a clinical result.
1023
1024 The reliability of an imaging agent reflects the reproducibility of the result (i.e., the value of a
1025 measure repeated in the same individual, repeated evaluations of the same image by different
1026 readers, or repeated evaluations of the same image by the same reader). (See the glossary for
1027 other related definitions.)
1028
1029 Many studies of imaging agents are designed to provide dichotomous, ordered, or categorical
1030 outcomes. We think it important that appropriate assumptions and statistical methods be applied
1031 in their analysis. Statistical tests for proportions and rates are commonly used for dichotomous
1032 outcomes, and methods based on ranks are often applied to ordinal data. We recommend that
1033 study outcomes be stratified in a natural way, such as by center or other subgroup category, and
1034 the Mantel-Haenszel[19] procedures provide effective ways to examine both binomial and ordinal
1035 data. We recommend that exact methods of analysis, based on conditional inference, be
1036 employed when necessary. We recommend that the use of model-based methods also be
1037 encouraged. These models include logistic regression models for binomial data and proportional
1038 odds models for ordinal data. Log-linear models can be used to evaluate nominal outcome
1039 variables.
1040
1041 In studies that compare images obtained after the administration of the test agent to images
1042 obtained before administration, dichotomous outcomes are often analyzed as matched pairs,
1043 where differences in treatment effects can be assessed using methods for correlated binomial
1044 outcomes. These studies, however, may be problematic because they often do not employ
1045 blinding and randomization. For active- and placebo-control studies, including dose-response
1046 studies, crossover designs can often be used to gain efficiency. We recommend that subjects be
1047 randomized to order of treatment. If subjects are not randomized to order of treatment, we
1048 otherwise recommend that the order in which images are evaluated be appropriately randomized.
1049 We recommend that study results from a crossover trial always be analyzed according to
1050 methods specifically designed for such trials.
1051

[19] For more on this topic, see Fleiss, Joseph, L., *Statistical Methods for Rates and Proportions*, 2nd ed., 1981, John Wiley and Sons, New York; and Woolson, Robert, *Statistical Methods for the Analysis of Biomedical Data*, 1987, John Wiley and Sons, New York.

1052 **B. Diagnostic Performance**
1053
1054 Diagnostic validity can be assessed in a number of ways. For example, both the unenhanced and
1055 enhanced images could be compared to the truth standard, and the sensitivity and specificity of
1056 the unenhanced image could be compared to that of the enhanced image. Two different active
1057 agents can be compared in the same manner. Diagnostic comparisons can also be made when
1058 there are more than two outcomes to the diagnostic test results. Common methods used to test
1059 for differences in diagnosis include the McNemar test and the Stuart Maxwell test.[20] In addition,
1060 we recommend that confidence intervals for sensitivity, specificity, and other measures be
1061 provided in the analyses. ROC analysis also may be useful in assessing the diagnostic
1062 performance of medical imaging agents over a range of threshold values.[21] For example, ROC
1063 analysis can be used to describe the relative diagnostic performance of two medical imaging
1064 agents if each test can be interpreted using several thresholds to define a positive (or negative)
1065 test result (see section IV.D.1). For all planned statistical analyses, we recommend that details
1066 of the analysis methods and specific hypotheses to be tested be stated prospectively in the
1067 protocol as part of the statistical analysis plan. We recommend that sponsors seek Agency
1068 comment on the design of and statistical approach to analyses before the protocols are finalized.
1069

[20] Ibid.

[21] For an introduction to this topic, see Metz, Charles E., *Basic Principles of ROC Analysis,* Seminars in Nuclear Medicine 1978;VIII(4):283-298. For a current treatment of statistical issues in diagnostic trials, see Zhou, Xiao-Hua, et al., *Statistical Methods in Diagnostic Medicine*, 2002, John Wiley and Sons, New York.

GLOSSARY

Note: Subjects in trials of medical imaging agents are often classified into one of four groups depending on (1) whether disease is present (often determined with a truth standard or *gold standard*) and (2) the results of the diagnostic test of interest (positive or negative). The following table identifies the variables that are used to estimate the parameters defined below.

Test Result:	**Disease: Present (+)**	**Absent (-)**	
Positive (+)	**TP (a)** true positive=TP	**FP (b)** false positive=FP	**m1 = a+b** = TP+FP total with positive test
Negative (-)	**FN (c)** false negative=FN	**TN (d)** true negative=TN	**m2 = c+d** = FN+TN total with negative test
	n1 = a+c = TP+FN total with disease	**n2 = b+d** = FP+TN total without disease	**N = a+b+c+d** = TP+FP+FN+TN total in study

Accuracy: (1) In common usage, *accuracy* is the quality of being true or correct. (2) As a measure of diagnostic performance, *accuracy* is a measure of how faithfully the information obtained using a medical imaging agent reflects reality or *truth* as measured by a truth standard or *gold standard*. Accuracy is the proportion of cases, considering both positive and negative test results, for which the test results are correct (i.e., concordant with the truth standard or *gold standard*). Accuracy = (a+d)/N = (TP+TN)/(TP+FP+FN+TN).

Comparator: An established test against which a proposed test is compared to evaluate the effectiveness of the proposed test. A comparator usually means an agent or modality approved for a similar indication. (See also the definition of *reference product*.)

Likelihood ratio: A measure that can be interpreted either as (a) the relative *odds* of a diagnosis, such as being diseased or nondiseased, for a given test result, or (b) the relative *probabilities* of a given test result in subjects with and without the disease. This latter interpretation is analogous to a relative risk or risk ratio.

1. For tests with dichotomous results (e.g., positive or negative test results), the likelihood ratio of a positive test result can be expressed as LR(+), and the likelihood of a negative test result can be expressed as LR(-). See the equations below:

$$LR(+) = \frac{\frac{a}{n1}}{\frac{b}{n2}} = \frac{sensitivity}{1 - specificity} = \frac{TruePositiveRate}{FalsePositiveRate} = \frac{\frac{a}{b}}{\frac{n1}{n2}} = \frac{PostTestOdds(+)}{PreTestOdds}$$

$$LR(-) = \frac{\frac{c}{n1}}{\frac{d}{n2}} = \frac{1 - sensitivity}{specificity} = \frac{FalseNegativeRate}{TrueNegativeRate} = \frac{\frac{c}{d}}{\frac{n1}{n2}} = \frac{PostTestOdds(-)}{PreTestOdds}$$

LR(+): *Interpreted as relative odds:* LR(+) is the post-test odds of the disease (among those with a positive test result) compared to the pretest odds of the disease.

Interpreted as relative probabilities: LR(+) is the probability of a positive test result in subjects with the disease compared to the probability of a positive test result in subjects without the disease.

LR(-): *Interpreted as relative odds*: LR(-) is the post-test odds of the disease (among those with a negative test result) compared to the pretest odds of the disease.

Interpreted as relative probabilities: LR(-) is the probability of a negative test result in subjects with the disease compared to the probability of a negative test result in subjects without the disease.

2. For tests with several levels of results, such as tests with results expressed on ordinal or continuous scales, the likelihood ratio can be used to compare the proportions of subjects with and without the disease at different levels of the test result. Alternatively, the likelihood ratio can be used to compare the post-test odds of disease at a particular level of test result compared with the pretest odds of disease. Thus, the generalized likelihood ratio can reflect diagnostic information at any level of the test result.

Negative predictive value: The probability that a subject does not have the disease when the test result is negative. Synonyms include *predictive value negative*. Negative predictive value = d/m2 = TN/(TN+FN).

By application of Bayes' Rule, the negative predictive value also can be defined as a function of pretest probability of disease (p), sensitivity, and specificity:

Negative predictive value = [(1-p) • specificity]/[(1-p) • specificity + p • (1-sensitivity)]

Odds: The probability that an event will occur compared to the probability that the event will not occur. Odds = (probability of the event)/(1 - probability of the event).

Positive predictive value: The probability that a subject has disease when the test result is positive. Synonyms include *predictive value positive*. Positive predictive value = a/m1 = TP/(TP+FP).

By application of Bayes' Rule, the positive predictive value also can be defined as a function of pretest probability of disease (p), sensitivity, and specificity:

1144
1145 Positive predictive value = (p • sensitivity)/[p • sensitivity + (1-p) • (1-specificity)]
1146
1147 **Post-test odds of disease:** The odds of disease in a subject after the diagnostic test results are
1148 known. Synonyms include *posterior odds of disease*. For subjects with a positive test result, the
1149 post-test odds of disease = a/b = TP/FP. For subjects with a negative test result, the post-test
1150 odds of disease = c/d = FN/TN. The following expression shows the general relationship
1151 between the post-test odds and the likelihood ratio: Post-test odds of disease = Pretest odds of
1152 disease x Likelihood ratio.
1153
1154 **Post-test probability of disease:** The probability of disease in a subject after the diagnostic test
1155 results are known. Synonyms include *posterior probability of disease*. For subjects with a
1156 positive test result, the post-test probability of disease = a/m1 = TP/(TP+FP). For subjects with a
1157 negative test result, the post-test probability of disease = c/m2 = FN/(TN+FN).
1158
1159 **Precision:** A measure of the reproducibility of a test, including reproducibility within and
1160 across doses, rates of administration, routes of administration, timings of imaging after product
1161 administration, instruments, instrument operators, patients, and image interpreters, and possibly
1162 other variables. Precision is usually expressed in terms of variability, using such measures as
1163 confidence intervals and/or standard deviations. Precise tests have relatively narrow confidence
1164 intervals (or relatively small standard deviations).
1165
1166 **Pretest odds of disease:** The odds of disease in a subject before doing a diagnostic test.
1167 Synonyms include *prior odds of disease.* Pretest odds of disease = n1/n2 = (TP+FN)/(TN+FP).
1168
1169 **Pretest probability of disease:** The probability of disease in a subject before doing a diagnostic
1170 test. Synonyms include *prevalence of disease* and *prior probability of disease*. Pretest
1171 probability of disease = n1/N = (TP+FN)/(TP+FP+FN+TN).
1172
1173 **Probability:** The likelihood of occurrence of an event, expressed as a number between 0 and 1
1174 (inclusive).
1175
1176 **Receiver operating characteristic (ROC) curve:** A graphical representation of pairs of values
1177 for *true positive rate* (or sensitivity) and the corresponding *false positive rate* (or 1-specificity)
1178 for a diagnostic test. Each pair is established by classifying the test result as *positive* when the
1179 test outcome equals or exceeds the value set by a given threshold, and *negative* when the test
1180 outcome is less than this threshold value. For example, if a five-point ordinal scale is used to
1181 rate the likelihood of malignancy for a tumor (e.g., definitely benign, probably benign,
1182 equivocal, probably malignant, definitely malignant), setting the threshold at *equivocal* will
1183 classify tumors as malignant (i.e., a *positive* test result) when the test outcome is at this level or
1184 higher and will classify tumors as nonmalignant (i.e., a *negative* test result) when the test
1185 outcome is less than this level. To generate an ROC curve, the sensitivity and specificity of the
1186 diagnostic test are calculated and graphed for several thresholds (e.g., all values of the rating
1187 scale). In a typical ROC curve, values for *true positive rate* (or sensitivity) are plotted on the
1188 vertical axis, and the corresponding values for *false positive rate* (or 1-specificity) are plotted on
1189 the horizontal axis.

1190
1191 **Reference product:** An FDA-approved drug product having an indication similar to that of an
1192 investigational drug or biological product to which it is being compared for the purpose of
1193 evaluating the effectiveness of the investigational drug or biological product.
1194
1195 **Sensitivity:** The probability that a test result is positive when the subject has the disease.
1196 Synonyms include *true positive rate*. Sensitivity = a/n1 = TP/(TP+FN).
1197
1198 **Specificity:** The probability that a test result is negative when the subject does not have the
1199 disease. Synonyms include *true negative rate*. Specificity = d/n2 = TN/(TN+FP).
1200
1201 **Truth standard (gold standard):** An independent method of measuring the same variable
1202 being measured by the investigational drug or biological product that is known or believed to
1203 give the *true* value of a measurement.