# Public Comment Form for QIBA Documents

Notes:

1. **Initials** identify the commenter to facilitate clarification of the issue and/or communication of the resolution.
2. **Priority**

**L**: Low. Typo or other minor correction that an editor can manage; requires no group discussion.

**M**: Medium issue or clarification. Requires discussion, but should not lead to long debate.

**H**: High. Important issue where there is a major issue to be resolved; requires discussion/debate.

1. **Line #** shows exactly where in the original document the issue occurs, and is necessary for sorting.
2. **Section #** shows in which section the issue occurs (e.g., 4.1.2)
3. **Issue**: Describe your issue; include enough to indicate what you see as a problem.
4. **Proposal**: Propose a resolution to your issue, e.g., suggested new wording or description of a way to address the issue; leave blank   
   if no resolution can be provided.

Stats: 167 Comments; 48 Low, 96 Medium, 23 High

Completed: 48 Low, 72 Medium, 12 High

Pending: 0 Low, 4 Medium, 7 High

Assigned: 0 Low, 7 Medium, 3 High

Remaining: 0 Low, 13 Medium, 1 High

Color Code: White – not done yet; Yellow – homework assigned; Green – homework done, review pending ; Grey – done

Useful Rules of Thumb:

(When reviewing a piece of text, generally a requirement…)

- Is it clear to an implementer and operator what they are expected to do here?

- It is clear how this would be tested?

- Would removing this requirement break the claim?

- Are we willing to fail someone for if they miss this requirement?

**Reviewed Document Filename:** **­­\_2011 07 28 Profile CT Advanced Disease V2 0f.pdf\_\_\_\_**

| # | Your Initial | Priority | Line # | Section # | Issue | Proposal | Resolution |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TSB | L | 44 | Closed Issue 3 | Dose material here is incomplete. | Reference the section where dose is addressed. Delete the second sentence in the closed issue. | Done |
|  | TSB | L | 51 | I | Don’t imply it has no use for faster change. | Re-word “Size quantification is helpful to evaluate tumor changes over the course of illness.” | Done |
|  | TSB | L | 54 | I | “this method” is ambiguous. | “this method” -> “RECIST” | Done |
|  | TSB | L | 55 | I | Wordy | Delete “depending on” | Done |
|  | TSB | L | 62 | I | Primary topic should standout | Break out last sentence into new paragraph;  “is expected to provide” -> “provides”;  “as well as” -> “and” | Done |
|  | TSB | L | 68 | I | staffs | staff | Done |
|  | TSB | L | 69 | I | PIs | Principle Investigators | Done |
|  | TSB | L | 70 | I | appropriate specifications | specifications | Done |
|  | TSB | L | 73 | I |  | Delete this line | Done |
|  | TSB | L | 80 | II | “their” | “lesion” | Done |
|  | TSB | L | 126 | III | Last sentence is a bit wordy and indirect. | “The profile does not intend to limit how equipment suppliers meet these requirements.” | Done |
|  | TSB | L | 134 | 1.1 | *“baseline”* scan | *“baseline scan”* | Done |
|  | TSB | L | 137 | 1.2 | “any timing” | “any other timing” | Done |
|  | TSB | L | 137 | 1.2 | Easier if separate topics are separated. | Start new paragraph at “Fasting” | Done |
|  | TSB | L | 148 | 1.4 | Avoid multiple usages for terms we want to be precise about. | “variance” -> “changes” | Done |
|  | TSB | L | 150 | 1.4 | “upper extremities” | “arms” | Done |
|  | TSB | L | 152 | 1.4 | “Feet first” | “Feet First by default” | Done |
|  | TSB | L | 155 | 1.4 | Subject Positioning:  “Feet first” | “Feet First” | Done |
|  | TSB | L | 158 | 1.5 | “inflates the lungs” See, this is a lung profile.  |  | Done |
|  | TSB | L | 173 | 2 | We use “participant”, “subject” and “patient” | Pick one and update consistently. | Kevin:  Committee prefers “Subject” – and note subjects may be patients. |
|  | TSB | L | 182 | 2 | Scan Plane is a separate topic from scan speed. | Break scan plane discussion into a new paragraph. | Done |
|  | GZ | L | 194 | 2 | Acquisition device to record anatomy | This is rather manual entry or what is the reference? | Kevin:  will “Just make sure it’s clear that the goal is to populate Anatomy in the image header. How it gets there might or might not involve manual entry.” |
|  | TSB | L | 194 | 2 | The 2nd and 3rd rows were updated to active voice, but the 4th-7th rows are still passive voice around the Acquisition Device. | “The Technologist shall…” | Done |
|  | TSB | L | 205 | 3 | “Spatial resolution is mostly determined by the scanner geometry” | “Maximum spatial resolution is mostly determined by the scanner geometry” | Done |
|  | TSB | L | 206 | 3 | “(which is somewhat under user control as the user usually gets to choose from a limited set of choices of reconstruction kernels provided at the scanner)” | “(over which the user has some choice)” | Done |
|  | TSB | L | 207 | 3 | “It is stated” | “Resolution is stated” | Done |
|  | TSB | L | 218 | 3 | “CT scanners” | “reconstruction algorithms” | Done |
|  | TSB | L | 224 | 3 | “not the same as resolution” | “not the same as spatial resolution” | Done |
|  | TSB | L | 227 | 3 | “they” | “it” | Done |
|  | TSB | L | 236 | 3 | “<Note that …>” | Break into new paragraph:  “Slice Thickness is “nominal” since the thickness is not technically the same at the middle at the edges.” | Done |
|  | TSB | L | 238 | 3 | “Characteristics need to be defined to optimize” | “Characteristics need to optimize” | Done |
|  | TSB | L | 241 | 3 | Redundant | Delete “currently” | Done |
|  | TSB | L | 244 | 3 | “For quantification of whole tumor volumes, the reconstruction software produces images that meet the following specifications” | “The reconstruction software shall produce images meeting the following specifications” | Done |
|  | TSB | L | 258 | 4 | Multiple Lesions:  This follows better after Common Lesion Selection | Swap this row with Lesion Volume Change. | Done |
|  | TSB | L | 258 | 4 | Recording:  Redundancy in first sentence. | Delete “shall be recorded” after the word metrics. | Done |
|  | TSB | L | 378 | Appendices | Give them each a letter to allow referencing them. | A: Acknowledgements  B: Background Information  C: Conventions and Definitions  D: Model-specific Instructions and Parameters | Kevin:  will do |
|  | TSB | L | 385 | App A | Easier if more visible | Make the last sentence a new paragraph. | Done |
|  | TSB | L | 450 | App A | “to determine tumor response (or progression) to treatment” | “to determine tumor response to treatment (or progression)” | Done |
|  | TSB | L | 468 | App B | “than reliance on linear tumor diameters” | “than linear tumor diameters” | Done |
|  | TSB | L | 469 | App B | “real clinical trial data” | “clinical trial data” | Done |
|  | TSB | L | 470 | App B | “volume measurements to be more reliable and often more sensitive to longitudinal changes in response than the use of diameters in RECIST” | Add commas for clarity.  “volume measurements to be more reliable, and often more sensitive to longitudinal changes in response, than the use of diameters in RECIST” | Done |
|  | TSB | L | 471 | App B | Since the prior discussion refers to methods that use one diameter and methods that use two diameters, the plural “diameters” here is mildly ambiguous. | “a uni-dimensional diameter” | Kevin:  will do |
|  | AZ | L | 475 | App B | Additional "accessing" in middle of sentence is a typo | Delete "accessing" | Done |
|  | TSB | L | 475 | App B | “to accessing assessing” | “to assessing” | duplicate |
|  | TSB | L | 480 | App B | “greater than 50% reduction in volume of tumor” doesn’t match the wording of the first part of the sentence. | “greater than 50% decrease in tumor volume” | Done |
|  | AZ | L | 482 | App B | "tumor shrink to a in a diameter" | should read "tumor shrink to a diameter" | Done |
|  | TSB | L | 482 | App B | “would result require that” | “would require that” | Done |
|  | TSB | L | 482 | App B | “shrink to a in a diameter” | “shrink to a diameter” | duplicate |
|  | TSB | M |  |  | Line numbers are most precise, but fall out of step when intermediate resolution drafts are reviewed, so Section # is helpful too. Encouraging people to randomly provide one or the other is the worst of both worlds for sorting/reviewing the comments. | Remove “(Please indicate either Line # or Section #)” from comment form. | Done |
|  | TSB | M | 3 |  | The title is a bit wordy for users and vendors to reference repeatedly. | Come up with an acronym (like IHE does) or a number (like ACRIN does) or something.  And shorten the title. | Anybody with a good idea  Moz: “QIBA Profile for CT Volumetry”  Neil: “CT Delta V”  Proposing  “CT Tumor Volumetry (CTV-1)” |
|  | TSB | M | 38 | Open Issues | Hard to reference specific open issues during comments. | Number the open issues, 1. 2. etc. | Kevin:  will do. |
|  | TSB | M | 38 | Open Issue 1 | The claim seems to represent the consensus of the committee on what might be reasonably achievable. It does not seem clear that the details currently specified in the profile are sufficient to reliably achieve the performance described in the claim in the field. | Consider inverting the claim to make the claim about the technical variation and the inference about the biological change.  Consider if we plan to do real-world validation of the claim and if so, how we want the surrounding text to read before and after completion of that validation groundwork. | Have reworded claim and are introducing concept of “levels of confidence” (See separate documents and process). |
|  | TSB | M | 38 | Open Issue 2 | There is a certain chicken-egg situation here so we need to get some systems working with the claim prior to being able to formally validate it. | Collect data about technical variation from sets of systems that comply with the profile operating on phantom and patient data. | With a stabilized specification here we can proceed to testing (as described in the levels of confidence) |
|  | TSB | M | 38 | Open Issue 3 | Somewhat case-by-case on whether to specify detailed results or detailed method. Have to evaluate each.  In some cases a fixed method will mostly guarantee the desired result and poor methods will not achieve the result and the method itself is not likely to be a useful point of innovation, in which case detailed method is the practical approach.  In other cases, we have reasons not to unnecessarily constrain the method (e.g. it IS an area of useful innovation, or it is an area of known variation which doesn’t impact our goals) so we describe the results, in which case it is important to be clear about the results and come up with good metrics (which is challenging, since unstated details are often sacrificed to meet the metric.) |  | Have made revisions to text to try to achieve appropriate balance. Still working on details of compliance testing. |
|  | TSB | M | 38 | Open Issue 4 | Seems reasonable to record the thinking of the group. | Add discussion of patients and lesions the profile is appropriate for. | Added discussion of assessable lesions and kept some requirements on subjects. |
|  | TSB | M | 38 | Open Issue 5 | 4cm/sec effectively excludes 16-slice and below systems and sets the cutoff at 64-slice. 64-slice is far from universal in the U.S. and even fewer abroad so this does preclude a lot of sites. | Consider exploring and addressing the discontinuity issues involved in “multi-breath” scans, either in this profile or in a separate profile. | Proposal generally accepted to resolve by permitting Total Collimation of 16mm (down from 20mm). The table speed requirement will continue to ensure single breath hold coverage and reconstruction requirements put appropriate bounds on acquisition parameters.  Most 16-slice scanners and above should be able to achieve this. Some examples that would meet this include:  (a) 16 x 1mm collimation with 0.5 second rotation time and pitch ³ 1.25 OR  (b) 16 x 1mm collimation with 0.4 second rotation time and pitch ³ 1 OR  (c) 16 x 1.25 mm collimation with 0.5 second rotation time and pitch ³ 1 OR  (d) 16 x 1.5mm collimation with 0.5 second rotation time and pitch ³ .833  Mike:  will consider if there are any issues with this solution. And describe the details/rationale between the table speed, total collimation, slice thickness, breath hold for the Discussion section.  Paul:  (Add instruction to relax diaphragm before scan while holding breath? Or delay 5 sec after inspiration before scan (for a fast scanner)(for a slow scanner, you’ll have your 5 seconds before you get down there?)? Full inspiration vs normal inspiration?)  (In general breath hold seems to be uniformly well understood/executed)  [Should we record if the patient achieved the breathold or not? Can the tech add this to the comments after the scan completes?][Useful work for future but beyond profile needs right now]  Neil – may need to consider in the context of non-lung? or is this overly constricting in order to meet lung |
| 1. u | TSB | M | 38 | Open Issue 7 | 5HU may be reasonable. For large masses it may be overkill. Depending on the contrast of the mass with background it may be necessary to shoot for 5HU.  HU “Accuracy” might also be an issue, but probably not a big one for volumetry unless segmentations are oversensitive. |  | Mike:  7. Q. Is 5HU StdDev a reasonable noise value for all organs? If it’s not, should we allow multivalued specifications for different organs/body regions?  Should we simply have several profiles?  A. Not sure I remember where the 5 HU standard deviation came from. We were using a standard deviation of 17HU in 1C project.  B. Here at UCLA, our Siemens Sensation 64 will yield a standard deviation of 17 HU for:  a. 120kVp, 50 eff. mAs, 1 mm thickness, B30F filter  C. To get this down to 5 HU would require:  a. Increasing the eff. mAs to 550, OR  b. Increasing the slice thickness to 2 mm AND increasing eff. mAs to 275  D. Let’s not have several profiles, but I suggest we change this value to at least 17 HU |
|  | TSB | M | 38 | Open Issue 9 | No. More compliance work is needed. | Especially describing requirements on techs, radiologists and measurement packages.  Review from the viewpoint of each actor thinking “What is the least I can do and still claim compliance” and see if that is likely to accomplish the goal and then fill any gaps. | Working on improving compliance. Work will continue. |
|  | TSB | M | 38 | Open Issue 9 | It is not clear who is subject to “compliance” and where they find their requirements. | Consider IHE style approach. Add explicit table of the actors in the profile. Explicitly indicate which actors are involved in each activity. | Kevin: done |
|  | TSB | M | 38 | Open Issue 10 | 15% is described as “technical variation”. So reader performance is considered part of technical variation? What are the other types of variation? | Clarify; perhaps in an Appendix that briefly defines the terms for measurement variability, the components and related statistical concepts? | Grace: |
|  | TSB | M | 49 | I | Mixing Roman numerals and Arabic numbers makes section references sometimes ambiguous. | Stick to Arabic. | Kevin:  will change to Arabic later.  Will need to coordinate with other groups to keep same across profiles in QIBA. |
|  | TSB | M | 56 | I | “major impact on patient management” and “has value” feels too deliberately vague. | Elaborate briefly on the value/impact. | Take the tack that “If you have decided to do volumetry, we’ll tell you a good way” but we won’t explicitly take a stance that volumetry is better than RECIST, or if it is better, what way it is better. |
|  | TSB | M | 58 | I | First two sentences are generic/QIBA process/FDA stuff that disrupts the flow of this CT profile. | Remove from summary or move elsewhere. | Kevin:  will delete mid-51 to mid-62.  Might want to move it elsewhere (Background) if it is not adequately covered there, and also perhaps in the Background note that the segmentation work here is even more broadly applicable than just volumes and may be built on later. |
|  | TSB | M | 72 | I | “Experts involved” is vague. Does this mean radiologists who make measurements, does this mean QIBA committee members, does this mean UPICT protocol proferrers? | Clarify | Neil:  will redraft to combine 69 & 72 and say:  “Clinical trial scientists designing end-points and protocols for the use of equipment for quantitative imaging”  “Scientists and physicians involved in the use of quantitative imaging in clinical practice”  Regulators |
|  | TSB | M | 80 | III | The title, summary and claim, etc address tumors. Scattered about are references to lesions. | Choose and fix. | Dan reviewed each tumor/lesion instance and determined if there is a distinction we need to preserve or not. |
|  | AZ | M | 87 | II | I don't understand why it follows from the previous statements that it means that technical variation is no more than 15%. Is "technical variation" the variation due to differences in acquisition? | Perhaps state that technical variation is expected to contribute no more than 15% to the measurement variability. | Claim text modified. Rationale added at end of B.2. |
|  | TSB | M | 96 | III | The opening sentence (and the figure title) are worded using academic paper language rather than specification language. | “The sequencing of the Activities specified in this Profile are shown in Figure 1.”  and consider moving the extensive figure title text out into a paragraph. Can probably shorten the text while doing that. | Kevin:  will break out the figure paragraph and shorten. |
|  | TSB | M | 97 | III | Since we never use the delta v notation elsewhere, we should just leave it out. | Drop delta v and delta TB notations. | Andy:  will drop deltaTB and the Interpret box from the diagram; retitle to change Tumor Burden to Target Lesions and leave deltaV in for now.  will look for places to leverage the deltaV. |
|  | TSB | M | 97 | III | “Directly process images to analyze change” | What is this allowing? | Kevin:  will go the route of the DICOM Grayscale pipeline. i.e. State the “Model” method and you can use other methods as long as your behavior appears equivalent.  Note that expressing as % growth seems to at least require an absolute volume measurement at time point 1. |
|  | TSB | M | 113 | III | The list should not “include the following”, the list should be complete. | Reformat as a definitive table.  Use exact actor names that will be used in the specification. | Kevin:  will do. |
|  | TSB | M | 118 | III | Activities are referred to elsewhere. Components are referred to here. And then Sections. | Replace “components” with “activities” | Kevin:  will do. |
|  | TSB | M | 119 | III | Not clear if this is normative | Use IHE style table. | Kevin:  will do. |
|  | TSB | M | 127 | III | This might be a good place to clarify the limits of the “requirements” on physicians. | Add paragraph “The requirements in this Profile do not codify a Standard of Care; they only provide guidance intended to achieve the stated Claim. Although doing so may invalidate the Profile Claims, the radiologist or supervising physician may deviate from the specifications in this Profile when required by the best interest of the patient or research subject.” | Neil:  will propose some wordsmithing.  Larry/Jim: will find an IRB person to review the language (or find boilerplate exculpatory language we can borrow?)  [Moz: The UPICT derived protocol will flush out any IRB issues if they exist. Close the Issue.]  [Does this text belong at this point in the document or higher up? Here seems good]  [Consider if we need to discuss a requirement on deviation from this? Perhaps not in the profile.]??? |
|  | TSB | M | 128 | III | Wording felt awkward | “This Profile is “lesion-oriented”. The profile requires that a given lesion be handled the same way each time. It does not require that lesion A be handled the same way as lesion B; for example lesions in different anatomic regions may be handled differently.” | Kevin:  will change to:  “This Profile is “lesion-oriented”. The profile requires that a given lesion be acquired and processed the same way each time. It does not require that lesion A be acquired and processed the same way as lesion B; for example lesions in different anatomic regions may be acquired or processed differently, or some lesions might be examined at one contrast phase and other lesions at another phase.” |
|  | TSB | M | 133 | 1.1 | At one point we decided to put all requirements in tables. This line has shall language outside. | Decide whether tables are used to wrap shalls or not, and make consistent. | Kevin:  Keep “shalls” in the Specification sub-sections (where the tables are) and out of the Discussion sub-sections, but otherwise don’t worry too much about keeping every shall inside tables. |
|  | TSB | M | 135 | 1.1 | “in no case more than the number of days before treatment specified in the protocol”  Q. which protocol are we referring to?  Q. is it for QIBA to say “in no case more”  Q. does this apply to clinical practice as well as clinical trials  Q. how many days offset warrants the dose of rescanning  Q. do we even need to address this since our claim is about biological change, not response to treatment | Discuss and revise. Consider dropping the phrase and generally defer to the clinical trial protocol for such details when the profile is used in the context of a clinical trial.  Alternatively, keep it and reference the specific details of the trial protocol used when this profile is used in the context of a clinical trial. | Moz:  The clinical trial specification has now been deleted as suggested. We agree that the content is related to the clinical protocol. Its inclusion reflected thinking before the scope of the profile became better understood. |
|  | TSB | M | 145 | 1.3 | Use of…: Shouldn’t the Radiologist be mentioned in the context of contrast? | Consider adding a Radiologist “shall” requirement | Neil:  will propose a sentence. |
|  | TSB | M | 145 | 1.3 | Image Header: Second sentence is describing alternative implementation methods. | Add a paragraph break and delete “that shall be”. | Kevin:  will delete second sentence |
|  | TSB | M | 151 | 1.4 | Do we need to record/duplicate the pillows? (placement, size, shape, material, amount of elevation…).  We mention it in the discussion then don’t address it in the specification. | Either describe how to address it or state that variations between scans are not expected to significantly impact the profile claim. | Gregory:  Added text in discussion and specifications. |
|  | AZ | M | 152 | 1.4 | CT scan in prone position is used more for assessing lung interstitium and functional lung changes. Not sure if prone position can be recommended as a scanning position for nodule / lesion volumetric assessment | QIBA team may consider the clinical setting and the scanning clinical practice at site and in oncology trials and recommend the most standard ones. | Gregory:  Proposal: Strengthen recommendation against prone position. Added to discussion. |
|  | TSB | M | 155 | 1.4 | We use variable phrasing:  “equivalent as used as baseline”  “same as for prior”  Etc. | Pick one code-phrase and use it consistently | Kevin: will do  “consistent with baseline” |
|  | TSB | M | 155 | 1.4 | If they have to do a scan that isn’t same as baseline (e.g. patient injury or equipment requires different positioning) should the subsequent scan return to baseline positioning if possible, or continue with the “new baseline”?  (Or any other parameter that has to change for some reason). | Clarify. OK if this is informative text for what to do to minimize effects (if we have anything useful to say).  Certainly want to record it clearly so people know it has happened. |  |
|  | TSB | M | 155 | 1.4 | Should address Centering since if not properly centered, dose modulation algorithm results will vary. | Add Patient Centering to spec and discussion. | Kevin: will do |
|  | TSB | M | 155 | 1.4 | Image Header:  “Table Height” | “Table Height and Subject Positioning” | Kevin: will do |
|  | TSB | M | 157 | 1.5 | Discussion section omits patient motion and the associated artifacts.  Later acquisition/reconstruction activities omit motion correction algorithms | Introduce mention of patient motion and the affect of such artifacts on quantitation.  Consider whether we should:  - provide guidance on when images become unusable for quantitation or need to be re-acquired? Or are we confident sites will make uniform judgements on this?  - permit/require motion correction algorithms in the acquisition/reconstruction activities |  |
|  | AZ | M | 158 | 1.5 | Breath holding is the technique used for abdominal assessment as well, not only for lung scan. | Would be clearer to include either a general statement on techniques used to reduce the motion artefact, either a more specific guidelines for scanning procedure, to include lung and abdominal-pelvic areas. | Review Proposed Text. |
|  | TSB | M | 163 | 1.5 | Image Header: second sentence should be reworded to be direct and specific.  (as mentioned in Open Issues, this needs to be addressed for most tables) | “The Acquisition Modality shall populate the following DICOM attributes in the Image:  Respiratory Motion Compensation Technique (0018,9170)  …” | Duplicate of #148 |
|  | AZ | M | 168 | 1.6 | It is not only the time-interval between the CM injection and scanning the parameter that influence the accuracy of lesion detection, but also the injection rate. | QIBA may want to mention that the contrast media injection rate / parameters should be the same use as at the baseline assessment. | Neil: will do |
|  | TSB | M | 171 | 2 | “CT scans for tumor volumetric analysis will be performed on equipment that complies…”  doesn’t quite make the point of the intent of the profile. | “In principle, CT scans for tumor volumetric analysis can be performed on any equipment that complies…” | Kevin: will do |
|  | TSB | M | 175 | 2 | In general we are putting the requirements in the Specification section and using the Discussion section to provide rationale, elaborate on terms, etc. | Expand the second last sentence into more of a rationale discussion and put it into a separate paragraph from discussion of using the same device. Incorporate the last sentence on FOV into the table. | Kevin: will do |
|  | TSB | M | 178 | 2 | The scout statement isn’t quite a rationale and isn’t quite a requirement. | Expand it into rationale (here) and/or requirements (in the table) or remove it. Break the “Pitch” discussion into a separate paragraph unless the scout rationale ties them together. | Discuss Scout: Pitch Done. |
|  | TSB | M | 179 | 2 | Are we permitting multiple breath hold as a valid implementation of the Profile? If so are there any compliance requirements? |  | Neil: “In some cases two or more breath-holds may be necessary to fully cover an anatomic requion. In such cases, different tumors may be measured on different breath-holds, but all the images used to measure each tumor shall be acquired within a single breath-hold to meet the claims of this profile.  If multiple breath-holds are required, the technologist shall obtain image sets with sufficient overlap to avoid gaps within the anatomic region.”  Comment-this second paragraph isn’t strictly required to meet our lesion based profile claim, but is important as an imaging protocol element if the scans are to be used to exclude new lesions |
|  | TSB | M | 182 | 2 | Does “Scan Plane may differ for some subjects” mean patient A will differ from patient B (which might be irrelevant for us), or patient A, scan 1 will differ from patient A, scan 2? And we should state the value of consistent scan plane. | Elaborate on the value of keeping it the same and clarify “differ”. | Mike:  Scan Plane (transaxial is preferred) may differ *between subjects* due to the need to position for physical deformities or external hardware. For an individual patient, consistency in positioning should be emphasized. |
|  | TSB | M | 184 | 2 | We say Total Collimation width might not be directly visible, but don’t tell them where to find it or derive it. | Elaborate. | Mike:  This is an introductory section and not a detailed user guide. Do we need to provide detailed instructions on where to find it or how to derive it? It would not be easy because this is done so differently across manufacturers. We could put a sentence in here to reflect this back to the manufacturers and say something like “Total Collimation Width (defined as the total nominal beam width) is often not directly visible in the scanner interface (see manufacturers reference material to determine this for a particular scanner make, model and operating mode). |
|  | TSB | M | 185 | 2 | We describe a tradeoff with Collimation but not our preference in the context of this profile | Indicate if we prefer to favor coverage, image quality or a middle balance. | Mike:  Again, this is an introduction and the particulars are very scanner dependent. Cone beam issues are much greater in those scanners with > 16 detector rows; but those same scanners provide greater coverage. Our preference is to get as much coverage as we can without introducing breathing artifacts or cone beam artifacts. So, how about a middle balance. “Wider Collimations …which may degrade image quality. Imaging protocols will seek to strike a balance to preserve image quality while providing sufficient coverage to keep acquisition times short.” |
|  | TSB | M | 189 | 2 | The 3rd sentence in the paragraph basically says the same thing as the 5th sentence. | Delete the 3rd sentence. | Kevin: will do |
|  | GZ | M | 194 | 2 | So far the profile leaves anatomic coverage open as ‘any’. In the specification is clear anatomic reference to thorax for scan duration. | Either stay with ‘any’ or add specification for other anatomies as well. | [Assume that the conspicuity requirements will make it clear to users of the profile which anatomy is not included. E.g. brain tumors will clearly not have sufficient conspicuity. Despite the selection of the acquisition parameters, it is expected that the segmentation algorithms will be able to handle the breadth.] |
|  | TSB | M | 194 | 2 | We require the tech to use settings that are the same as prior but don’t provide a mechanism for achieving that. | Consider requiring modality to be able to query/retrieve prior/baseline and present certain values to the tech.  Once DICOM Sup 121 (Protocol Storage) is complete, might want to mention it as well. | Maybe in the future. Add a note similar to Analysis. |
|  | TSB | M | 194 | 2 | Tube Potential:  Different models sometimes have different available kVp settings. Exactly the same value may not be possible (or appropriate based on model design) | Provide guidance on the use of different values for different models. | Mike:  ”The Acquisition Devise shall be set up so as to achieve the same kVp (or as similar as possible) for fall scans” |
|  | TSB | M | 194 | 2 | Image Header:  There should be a direct correspondence between items in this list and the parameters above however there are frequent mismatches/gaps. | Consider removing the Image Header row and distributing it into each of the parameter rows so the pairing is visible and easily maintained. E.g. for Scan Plane  “The Technologist shall set the scan plane to be the same as for prior scans.  The Acquisition Device shall record the actual Gantry/Detector Tilt (0018,1120). “ | Kevin: Conflicted with simplicity of Comment 115. Leaving it in the row at the bottom.  (Consider new column?) |
|  | TSB | M | 194 | 2 | Single Collimation Width:  Exact meaning of the term (re: detector structure, adjustable collimation, etc) was not clear. | Clarify definition in the Discussion section. | Mike:  In IEC standards, “T is the nominal tomographic section thickness” It is this value that we want to be <= 1.5 mm. So, perhaps the table row should say “Nominal Tomographic Section (T)” and the next column would say ”The Acquisition Device shall be set up so as to achieve a Nominal Tomographic Section (T) <= 1.5 mm.” |
|  | TSB | M | 194 | 2 | Does a discussion of metal artifacts go here, or in patient preparation (do they have metal in places that might cause acquisition artifacts that impact quantitation), or in reconstruction (can we minimize?), or in image analysis? | At least raise a flag in one of the locations. |  |
|  | TSB | M | 195 | 3 | Can’t reference the informative material and the normative material separately. | Consider giving Discussion and Specification their own subsections (e.g. 3.1 and 3.2)  (And repeat pattern for other Activities) | Kevin: will do example |
|  | TSB | M | 197 | 3 | The first paragraph misses a couple points. | “Image reconstruction is modeled as a separate Activity in the QIBA Profile. Although it is closely related to image acquisition, and is usually performed on the Acquisition Device, reconstruction may be performed, or re-performed, separate from the acquisition. Many reconstruction parameters will be influenced or constrained by related acquisition parameters.” | Kevin: will do |
|  | TSB | M | 202 | 3 | Should clarify the nature of our noise concern.  (And is this a practical concern? How noise-sensitive are current algorithms compared to the radiologist reading?) | “increase in noise.” -> “increase in noise which may degrade segmentation and quantification of tumors.” | Kevin: will do |
|  | TSB | M | 208 | 3 | We still have two definitions of resolution. If one is numerically computable from the other, we don’t need both. If they are subtly different, then we shouldn’t use them equivalently. | Choose only one definition, or else clarify the distinction between them. Could also make one normative and the other informative | Mike:  Removed FWHM |
|  | TSB | M | 213 | 3 | Not clear whether “using thicker slices” refers to acquisition parameters or reconstruction parameters or both. | Clarify | Mike:  “Noise (pixel standard deviation) can be reduced by reconstructing images with greater thickness for a given mAs. A constant value for the noise metric might be achieved by increasing mAs for thinner reconstructed images and reducing mAs for thicker reconstructed images. |
|  | TSB | M | 214 | 3 | It’s not clear what this paragraph is trying to tell the radiologist/technologist/device developer. | Clarify or drop. | Mike:  = not sure what to drop here.  Perhaps we could change the sentences to “The standard deviation does have limitations since it can vary with different reconstruction kernel, which will also affect spatial resolution. While the Noise Power Spectrum would be a more comprehensive metric, it is not practical to calculate (and interpret) at this time. Therefore, the standard deviation metric is currently the preferred measure. |
|  | TSB | M | 224 | 3 | The term “inherent resolution” is introduced but not sufficiently clarified. Is it the resolution inherent in the acquisition hardware? It is different than how the data is reconstructed yet it is strongly affected by the kernel. | Clarify | Nick  Pixel size in each dimension is not the same as spatial resolution in each dimension. The spatial resolution of the reconstructed image depends on a number of additional factors including a strong dependence on the reconstruction kernel. |
|  | TSB | M | 225 | 3 | The term “data field” is introduced but not sufficiently clarified. | Clarify or use a different term (we use “images” elsewhere but maybe we need a term for the image stack/volume). | Kevin:  changed to dataset like following paragraph. |
|  | TSB | M | 226 | 3 | Seems to imply that it’s OK to scan at different resolution as long as it’s higher resolution. Wouldn’t that also potentially change the segmentation/quantification? | Clarify the requirement/policy/guideline. | Nick:  Remove this sentence since matching resolution isn’t part of the specifications (we only include minimum requirements). If the goal of the sentence is to make a statement about the potential for improved volumetric consistency using the same spatial resolution then a different sentence should be included.  <Added requirement for consistent resolution> |
|  | TSB | M | 238 | 3 | What are we trying to ask for here? This seems to imply a kernel be individually optimized for each lesion which is not our intent (if only for practical reasons). | A different Well-suited kernel for each body region?  A single Compromise kernel for entire scan? | Neil:  “The choice of reconstruction kernel influences the appearance of tumors on the reconstructed images, and may influence image measuremtns. A softer kernel can reduce noise at the expense of spatial resolution. An enhancing kernel can improve resolving power at the expense of increased noise. Images shall be reconstructed using one or more reconstruction kernels suitable for the anatomic region and tissue imaged (e.g. lung, soft tissue). A tumor shall be measured on images reconstructed using consistent kernels at each time point.” |
|  | TSB | M | 245 | 3 | The “same as baseline” requirements are missing. | Add them similar to Acq. Table | Kevin: will do |
|  | TSB | M | 245 | 3 | Are we going to allow iterative reconstruction and how would it affect these parameters? Can the baseline be iterative and the follow-up not, or visa-versa? | Explicitly address. | Ehsan:  <Consider allowing IR algorithms but mandating “same” type of algorithm be used as baseline? At least until we have a more confidence that it doesn’t effect repeatability?>  <So how should we revise the sentence “The effects of iterative reconstructions on quantitative accuracy and reproducibility are not fully understood as of this writing of this profile version.”?>  <Maybe something along the lines of “as long as the noise and resolution specs are met, it seems like IR should perform within our requirements>  Since we don’t have information on it’s reproducibility and impact on accuracy the safest thing is to not permit it in the profile for now. |
|  | TSB | M | 245 | 3 | The requirements are passive requirements referencing the Reconstruction Software. | Make active requirements on Tech.  Consider deleting all the shall language in the table in deference to the shall at the top of the table (although this would make it harder to weave in the header requirements in each row). | Kevin: will do |
|  | TSB | M | 245 | 3 | Spatial Resolution:  We don’t describe how Axial FWHM is measured. Thin Wire? Bead? Are there standard protocols for this? | Resolve. | Mike:  Dropped the spec. |
|  | TSB | M | 245 | 3 | Reconstruction Overlap:  Wouldn’t =0 also be no gap? | Change “>” to “>=” | Kevin: will do |
|  | TSB | M | 245 | 3 | Voxel Noise:  This may be hard to achieve at low dose, but perhaps that’s the point. |  | Agreed. Noise target revised from 5HU to 18HU. |
|  | TSB | M | 245 | 3 | Voxel Noise:  Are we dictating that it may only be measured with a 20cm phantom? Would a different size be better? Should match the std FOV of the acquisition device. | Clarify/advise | Mike:  Considered allowing “equivalent” values/materials/size with an obligation on the tester to prove they are equivalent. Finally decided not to allow for this.  Voxel Noise: The Reconstruction Software shall be set up so as to achieve voxel noise standard deviation of < 18 HU in a 20 cm water phantom. Equivalent performance in alternate sized phantoms should be allowed, but the burden of demonstrating that the performance in an alternate phantom is equivalent to this specification shall fall to the group utilizing the alternate phantom. “  In other words, this is our standard. If you want to use a different sized phantom, go ahead , but YOU have to prove that what you are doing is equivalent. Not sure what else to offer here. There is NO WAY we can determine what the equivalent specs are for all of the different possible phantoms. |
|  | TSB | M | 245 | 3 | Recon FOV: Is it our intent to prohibit “trimming the fat”? I.e. we are dictating that sites shall sacrifice resolution in order to cover all tissue outside the lungs. | Confirm. | Neil:  Allow trimming, but require consistency. |
|  | TSB | M | 245 | 3 | Slice Thickness:  The acquired slice thickness constrains what thicknesses you can easily reconstruct.  If scanned on different devices, they may have different acq slice thicknesses available. What is our requirement/guidance?  E.g. if 2.5 at baseline is 2.0 OK for followup? Visa versa? | Clarify/advise.  In the case where the baseline is 2.5, and the current acquisition data is 2mm, should the reconstruction bend the “consistent with baseline” and reconstruct 2mm, or should it process for 2.5mm? | Nick  "Slice thickness consistent with baseline" could be defined as within 0.5 mm of baseline subject to 2.5 mm maximum. |
|  | TSB | M | 245 | 3 | Reconstruction Kernel Characteristics:  The spatial resolution and noise characteristic requirements go a long way toward constraining the kernel appropriately. Do we need a requirement around HU values/ramp?  Is there a potential problem if segmentation pays too much attention to HU Thresholds? | Clarify/advise | Nick  “Kernel consistent with baseline” could be defined as utilizing the kernel most closely matching the kernel response of the baseline (the same kernel if imaging on the same system). This should reduce the chance the spatial resolution/noise specifications are met but the kernels produce very different textured images.    We may want to also include a specification that the reconstruction type must be consistent with baseline (IR can’t be mixed with FBP).  <And seems unlikely the exact HU will vary too far, and unlikely the segmentation will be TOO sensitive to it. There is some potential for an issue here. 1C saw some negative bias on spiculated lesions potentially due to this> |
|  | TSB | M | 248 | 4 | Break up the topics. Address alternate methods more clearly. | Each lesion is characterized by it’s volume change relative to prior image sets.  This is typically done by determining the boundary of the lesion (referred to as segmentation), computing the volume of the segmented lesion, and taking the difference of the lesion volume for the current and prior.  Segmentation may be performed automatically by a software algorithm, manually by a human observer, or semi-automatically by an algorithm with human guidance/intervention, for example to identify a starting seed point, stroke, or region, or to edit boundaries.  Methods that calculate volume changes directly without calculating volumes at individual time points are acceptable so long as the results are compliant with the specifications in this profile. | Kevin:  will do. |
|  | TSB | M | 248 | 4 | While we’re recognizing different approaches, should allow for surface smoothing, partial volume, etc. | “Volume Calculation from a segmentation may or may not correspond to the total of all the segmented voxels. The algorithm may consider partial volumes, do surface smoothing, tumor or organ modeling, or interpolation of user sculpting of the volume.” | Kevin:  will do. |
|  | TSB | M | 248 | 4 | What about additional pre-processing of images by the Image Analysis Tool? | Indicate that additional pre-processing of images by the Image Analysis Tool is permitted. | Kevin:  will do. |
|  | TSB | M | 258 | 4 | Common Lesion Selection:  We introduce the term “common set of lesions” but don’t define it.  We refer to lesions being measured by “all readers” but don’t clarify who that is in a clinical practice context and a clinical trial context. | Re-draft sentence. (or two)  And consider more clearly distinguishing the task of matching a lesion to a prior image/instance of the same lesion, from the requirements for handling multiple lesions (e.g. create an overall tumor burden metric? provide lists of tumor results? etc) | Gregory:  Make distinction between labeling each lesion for comparison across timepoints and labeling a set of lesions for assessment by more than one reader, when applicable. |
|  | TSB | M | 258 | 4 | Lesion Volume Change:  “measure lesion volume change with variability less than +/-15%” | Clarify shall measure each lesion.  Describe how individual measurements relate to the aggregate performance requirement. | David:  address together with Comment #163  Done – see #163 – and the claim is 30% not 15% |
|  | TSB | M | 258 | 4 | What is the “granularity” of the performance scores? We could have a test set and score the algorithm on the match between it’s total tumor burden estimate and the “actual”. We could total up the errors for each individual tumor. We could map and characterize the distribution of errors. Our metric should reflect what we are trying to achieve. This is not always obvious/easy. | Discuss and decide. Will probably want to revisit as we get more experience. | David:  address together with Comment #163  Done - see #163 – we should not total, IMHO, but rather we should quote performance on individual lesions over a specified number of repeats for a specified number of lesions. |
|  | TSB | M | 258 | 4 | Lesion Volume Change:  Performance can be affected by the accuracy or variations in the seed point or axis provided. | Consider preparing the test set with test “inputs” (either with a “click here” dot on the image, or some method for feeding coordinates to the application). | Added to test set discussion. |
|  | TSB | M | 258 | 4 | Lesion Volume Change:  Some algorithms may use landmark structures, which may be adversely affected by congenital or surgical abnormalities. | Consider whether we should flag this issue in the profile. Probably not. Can we come up with a rule of thumb for what gets discussed in the Profile and what is left to individual product documentation? | Not likely to be a significant issue. |
|  | TSB | M | 258 | 4 | Recording:  Should we add a requirement that if the same Analysis Tool is used for processing the followup images, that it should use the same parameters as used for the baseline? If so, we should require that it record the parameters used for each analysis in addition to recording those used for compliance testing.  Of course, some of the parameters (denoising settings) might be particular to the image set and shouldn’t be copied, while others might be shape oriented. | Discuss. | Added note recommending that a single package analyze both current and prior time points in order to get consistent processing. |
|  | TSB | M | 258 | 4 | Recording:  Second sentence is overloaded and a bit confusing. What does “(and reload for review from) region specification” mean? | Clarify and break up. | Kevin:  modified and requirement dropped. |
|  | TSB | M | 258 | 4 | Recording:  “one or more of the following output formats” is asking for lack of interoperability. | Make one mandatory (both read and write) to enable interoperability. | Kevin:  will do. |
|  | TSB | M | 258 | 4 | Recording:  Do we need to record the software version that was used for this processing?  Do we need to record all software components in the system?  Do we need to record hardware components?  E.g. Changing a low level numerical processing library or a different GPU could result in a slight change in precision for a sqrt() which could give a different result, particularly for “chaotic” cases.  In one case, Microsoft changed a sin() function to fix a security hole (so the changes may be mandatory and/or not recognized as affecting segmentation) and it changed some behaviors. | Figure out what our goal is (e.g. identically replicate a result VS flag potential sources of variance to the extent that is practical) and derive a reasonable cutoff point. | Kevin:  <Assume a medical device is a static stack and then version the stack>  <Should record the operator(s) involved as well, as good policy if not for meeting the claim. Whether it is an actual identity, or just an ID to correlate with prior measurements is fine. Having it would be important for analyzing operator as a factor, but maybe analysis is outside the scope of the profile which is about performance.> |
|  | TSB | M | 259 | IV | Compliance requirements already exist in the form of the Profile “shall language”. As it exists now, Section IV is mostly redundant for actor requirements. If we want a section on site validation of the entire chain, that would be useful, but I don’t know if I’d call the section Compliance in that case. | Revise or remove. | Kevin:  Propose revised text.  Issue of maintaining parallelism between domains.  Consider keeping Compliance as the title and then make Performance Assessment a component of that. The other component being to comply with section III.  Discuss with other domains. |
|  | TSB | M | 290 | References | Do we still need all of these? |  | Andy:  Will update to convert these to embedded references/links once the document is stable |
|  | TSB | M | 441 | App A | Current wording implies that “deployment in real-world settings” does not constitute use. | Come up with another term for “released for use”. | Andy  Delete the sentence. |
|  | TSB | M | 444 | App A | Sentence uses both the term Conformance Statement and Compliance Statement. Let’s pick one. | Conformance Statement. | Kevin: will do |
|  | TSB | M | 445 | App A | “Conformance Statements represented as an appendix called Model-specific Parameters”.  This is incorrect. | Either delete the “represented as…” phrase or add an appendix showing the format of Conformance Statements (like IHE) and review the text at the beginning of Appendix D and see if it needs clarification. | Kevin: will clarify  <add a requirement for Appendix D tables for one (or more) reference patients as part of the conformance statement>  <Discuss/propose with other QIBA committees>  Mike:  <Will provide some text on what the table “means”>”Reasonable” but not recommended or optimized  “These are reasonable protocols that are expected to meet the profile claim…. language about size.”  Not sure how to proceed here. What we want are examples of protocol charts in which technical parameters are spelled out for specific makes and models of scanners that will result in compliance with the QIBA profile. What should we call this? If it is not a Conformance statement, then what is it?  We can also provide some wording like “A protocol compliant with the QIBA profile will include some description as to how the protocol should be adjusted appropriately for patient size” and we can leave it at that. This is because we want them to have some adjustment for patient size, but there is no agreed upon single method as to how to do that. (but it is better than NOT having them adjust for patient size). |
|  | TSB | M | 478 | App B | “Secondly, changes in volume **are** less subject to either reader error or inter-scan variations.”  Seems more certain that “potentially less prone” in line 484. | Decide which we like and change to be consistent. | Andy:  Revise wording to “are potentially” or insert references substantiating “are” and consider strengthening 484. |
|  | AZ | M | 482 | App B | "diameter of less than 158mm" - the tumour was only 21mm to start with, so cannot shrink to 158mm | 158mm should read 14.7mm (please double check) | Andy:  will check/fix. |
|  | GK | M | 482 | App B | “response would result require that the tumor shrink to a in a diameter of less than 158 mm, but would correspond to a decrease in volume all the way down to 17702145 mm3” | I believe 158 mm needs to be close to 14.7 mm | duplicate |
|  | TSB | M | 482 | App B | 158mm and 17702145mm are clearly not the correct values | Fix. | duplicate |
|  | AZ | M | 483 | App B | a tumour volume of 17702145mm3 doesn't appear to be correct | 17702145mm3 should read 1663mm3 (please double check) | Andy:  will check/fix. |
|  | TSB | H | 3 |  | Although it would be nice, it seems overly ambitious to assume that one acquisition protocol can be specified that is appropriate for any solid tumor of any type of cancer anywhere in the body. For example, the profile requirements do not seem well tuned to a patient who has presented with a brain tumor.  More specifically, the literature in Table B-1 is almost entirely focused on lung. | Clearly constrain the profile scope in the Title, Summary and Claim to be lung nodules and allow for addressing associated metastases but with words about the potential sub-optimality for the metastases.  In the future, other profiles can be developed that re-use parts of this one as appropriate. | Moz and Larry:  The key point is that the margins be conspicuous and the geometry is simple enough.  (Conspicuity requirements would disallow using this for primary brain tumors. Although they can be measured, we have no concrete evidence our profile will perform up to the claim for those.)  In terms of formatting, it may make sense to change “measureable” in the claim to “appropriate” and then define appropriate in a section under the claim.  [Do we have to rule out spiculated tumors since they are more susceptible to fine partial volumes? No. Geometrically simple allows for spiculated to be included]  David: I sympathize with the comment by TSB – there really isn’t sufficient evidence to substantiate the claim outside the lung, but on the other hand, that is the point of having this profile, to use for advanced metastatic disease drug trials, which involve the whole body, including brain metastases (as opposed to primary brain tumors like GBM), although one would not normally measure these on CT as opposed to MRI, or measure them at all (treat as non-target or new disease, not measurable disease) – perhaps this can be addressed by creating the appropriate nominal test data set for compliance testing with a sufficiently broad range of lesions; see also #150. |
|  | AZ | H | 38 | Open Issue 1 | Quite hard to see what the absolute gain would be in terms of accuracy of measurement over standard RECIST 1.1 contrast enhanced 5mm slice width scans - has this work been done or is it planned within QIBA. | QIBA team could consider what are absolute requirements & what are should haves in terms of references to experiments which indicated the impact on accuracy and precision of volume measurements at >=10mm relevant to oncology.  For instance a lot of the references look to be pulmonary nodules in screening application where slice width <=2.5 would seem more relevant than in an oncology treatment setting where lesions >= 10mm would be the starting point.  None of the references pointed to work done at the clinical standard of 5mm for CT scans to assess tumour response assessment rather than thin slices for screening. | Grace:  will write up a rationale and the implication of the 30% number (It is a good number, it gives you some power in study design, but we should not make both the 15% claim and the 30% change confidence claim, since there is not a direct equivalence between the two.)  Neil:  will propose a resolution to underlying question A:  Are we making this so strict that it wouldn’t be accepted in clinical practice?  The profile could point out that the radiologist can read the thick slices and even place a seed on the thick slice which the software could translate to the thin slices and do it’s measurements on the thin slices.  We should stick to 2.5mm slices for practical purposes and if we have to grow the variance (change our claim) do that. It needs to be practical. We can stretch in a future profile for thinner slices. There are nasty lesions out there that readers are struggling with.  Given the profile requirements on the (larger) size, conspicuity and shape of the tumors, it seems likely that 30% will still hold even if we go up to 2.5mm slices.  Mike:  will propose a resolution to underlying question B:  Are we asking for higher dose than clinical practice in order to get less noise in the thin slice?  No. With the increase to 18HU, we are not pushing dose for 2.5mm slices.  Observations:  - on high contrast lesions, dose affects the variance but not the bias of measurements  - some Clinicians are not going to measure lesions they won’t be able to repeatably measure (e.g. because they expect them to merge, or because the shape is odd and hard to repeat)  - tumors <10mm are “non-target” and are followed visually in clinical trials rather than quantified  - RECIST isn’t a fully qualified biomarker. It’s use-dependent. Qualifying specific uses is easier than qualifying open ended/general use |
|  | TSB | H | 38 | Open Issue 6 | Good question. Pixel Std Dev is typical, but it is a poor metric for non-linear reconstruction methods. | Investigate Noise Power Spectrum. This provides more nuanced data (which in turn means, unfortunately, that it does not lend itself well to a simple cutoff metric). | Mike:  will consider and propose a conclusion. |
|  | TSB | H | 38 | Open Issue 8 | Agree with Answer. | Explicitly specify the name and tag of each necessary DICOM field. This needs to be done; gaps identified and fixed. | Kevin: will do later when stable (large task). E.g.  “The Acquisition Modality shall populate the following DICOM attributes in the Image:  Respiratory Motion Compensation Technique (0018,9170)  …” |
|  | TSB | H | 38 | Open Issue 9 | It may be useful to define methods to validate/measure the performance of each activity independently (useful for actor compliance testing and for doing QA on the chain) and also have a validation for the entire chain (useful for sites validating their particular installation) | Consider actor, activity and full-chain validation methods/metrics. | Moving in that direction. WIP. |
|  | AZ | H | 59 | I | Is it intended that this protocol and claim be applicable to the measurement of change in any solid tumour by CT, or does it just apply to pulmonary nodules? All the references in the table in the Background Information section are for pulmonary nodules. | Please clarify in the Executive Summary and throughout the document whether the profile refers to pulmonary or all nodules. If it applies for all tumours, please provide supporting references in table in Background Information section. | Larry and Andy:  will identify references (e.g. from the briefing document) to support our position that measurement variability outside the lung will also be within our claim, and the variability data from our groundwork is equally applicable to variability outside of the lung.  May need a paragraph describing the differences and extrapolation of findings involved. Lung is easier, but there are some reader studies for liver and lymph.  <Kevin found the submission. The full submission is too large to include here but could go in Annex B. > |
|  | AZ | H | 83 | II | Claim is addressed in percentage change and profile in absolute volume change | adjust profile to report percentage change in line with the claim | Will keep profile in terms of % (not change in mm3) |
|  | GZ | H | 89 | II | Volume variability of 30% | It would be beneficial to have a justification for 15% measurement variability that is referenced from 2D measurements; Is there a difference in variability in x,y and z? | Mike and Grace:  Not clear that difference x/y/z variability have a substantial affect on our profile claim  (15/30 question duplicates #146) |
|  | AZ | H | 142 | 1.3 | The use of contrast media in assessing tumor boundaries and ultimately the change during the treatment is critical. A non-contrast CT will not permit an accurate characterization of the malignant visceral/nodal/soft-tissue lesions. | QIBA team could consider that the use of contrast media should be mandatory in the context of oncology and for assessing tumor response. | Neil:  will draft proposed text based on the discussion conclusions.  Specify must use contrast, same phase, etc. Probably can’t specify detailed contrast protocols due to local variations in practice, but can provide basic requirements and require consistency.  (Consider re-emphasizing that sites may deviate, they can still measure, etc. just that they’ve left the scope of the claim.)  (Chest only might not need this, but since we’ve set the scope of the profile to metastatic, we’ll have to deal with non-contrast in a Chest only Profile)  Observations:  - some lesions will not be conspicuous enough to be measured to our spec without contrast. Contrast is good practice – lesions get missed on non-contrast  - alternative was to limit the scope of the profile (i.e. expand requirements for conspicuity and tumors that do not meet conspicuity or appear later and lack conspicuity are outside the profile.) |
|  | SF | H | 145 | 1.3 | Not all power injectors interface with the CT so DICOM headers may not contain the info. |  | Neil:  will propose a list of the minimum set of specific details that must be recorded for adequate repeatability and analysis |
|  | SF | H | 145 | 1.3 | Contrast will be used as it was at baseline for all subsequent scans. This will be problematic because of changes in Nephric health status. | Decide if unenhanced alternative is okay, such as non contrast MRI of abdomen and pelvis. Determine if it would be acceptable to continue without IV contrast for chest only. | Kevin:  will add text to say:  “True, but that just means that patient has left the protocol.”  May need clarification that deviating from the protocol is up to the physician. The consequence is being outside the scope of the claim. You can still try and do what you can.  (merged into #153) |
|  | TSB | H | 155 | 1.4 | The baseline/first contact with the patient is the most important opportunity to optimize the characteristics of the scan since all subsequent scans should strive to duplicate this, and if the baseline is sub-optimal, it puts the follow-ups in a quandary. | Discuss and clarify. Informative? Normative? | Kevin:  will add material emphasizing that the baseline is important.  Change references to “prior scans” to “baseline”.  Clarify that breaking a “shall” is a protocol deviation. Deviation loses your QIBA Claim. How study sponsors and others decide to deal with deviations for their own purposes. |
|  | SF | H | 168 | 1.6 | The acquisition device shall record actual timing and triggers in the image header | This is not available to my knowledge | Kevin:  DICOM supports:  Contrast/Bolus Start Time  Acquisition Time  Contrast/Bolus Stop Time  Contrast/Bolus Total Dose  Contrast Flow Rate(s)  Contrast Flow Duration(s) |
|  | SF | H | 187 | 2.0 | Slice width affects Voxel size along Z axis. Smaller voxels are preferable to reduce partial volume effects and provide higher accuracy due to higher resolution. This is all well and good but at the techniques that are suggested the images will be degraded. | Increase technique. | Mike:  <will draft a note that the technique will drive the noise which is context/region sensitive, and that there are noise requirements later in the profile. And we don’t expect the same technique for lung and abdomen.  In high contrast work, like lung nodules, the measurement may not be as susceptible to noise as the visual perception is. Below the diaphragm it does become an issue for measurement as well though.>  ??? We have allowed image thickness up to 2.5 mm .??  Based on previous discussions, this thickness and these techniques will work for many lesions (10 mm, and tumor margins are visible). I think we should be ok. |
|  | SF | H | 194 | 2.0 | “The acquisition device shall be capable of performing the required scans at an axial rate of at least 4cm per second”? | I would think that most facilities would be scanning chests in helical mode. | Mike:  <will confirm the appropriate AAPM term, but yes, helical is expected. Will rename axial rate to… table speed (or it’s equivalent rate of coverage of the anatomy)  May need a variant/exception to allow for systems that can do full coverage in a single rotation.>  “The acquisition device shall be capable of performing the required scans at a table speed of at least 4cm per second”?  May need a variant/exception to allow for systems that can do full coverage in a single rotation. However, no scanner currently used in diagnostic imaging can cover the entire chest or abdomen or pelvis in a single rotation. |
|  | GZ | H | 245 | 3 | Slice thickness and profile claim | Is there any literature justification to provide evidence that for slice thickness <2.5 mm there is variability of 30%; Literature review is mainly referring to 1.25mm | Larry & Binsheng:  will identify literature that supports 2.5mm 30% and draft brief text for inclusion in Appendix B in the profile.  (consider rationale that the references support 1.25mm but do so for smaller nodules than our 10mm, so we can infer that we could still reach 30% on 10mm+ nodules using 2.5mm slices)  (literature may include our groundwork which has papers almost ready to go out for submission)  Observations:  - dropping the protocol to 1.25mm would lose most of our constituency |
|  | TSB | H | 248 | 4 | The current requirements are all on the Analysis Tool, implying no responsibility on the operator (Radiologist?). Most current tools are put forth with the understanding that the human is the ultimate arbiter of the quality of the results and either modifies or verifies/approves each of the segments produced. | Address the verification role/responsibility of the operator more explicitly in the discussion and consider whether there are some requirements to add. E.g. should the software record the identity of the operator who is responsible for the results. | Discussion of the effect of the Analysis Tool operator/approver, and the need to record their identity added. Also added discussion of tumor selection.  Comment: although full FDA Part 11 compliance is important for requlatory trials, it is probably beyond the scope of this profile.  Comment-selection of the imaging and contrast protocol is typically at imaging site and does not belong in this section. I think we have covered this elsewhere. |
|  | TSB | H | 258 | 4 | Multiple Lesions:  We require a lesion identifier for correlation across time points, but don’t describe or require an interoperable mechanism to make it feasible. | Consider DICOM SR tools. | Kevin:  will add text to the effect that the measurement software is responsible for matching the current volume measurement with the prior measurement for the same tumor to get the change. Since we have not provided a method to ID and match them, this means the software must generally be prepared to measure both the current and the prior in the same session. |
|  | TSB | H | 258 | 4 | Lesion Volume Change:  We specify “variability less than 15%” but don’t describe or require a validation procedure. | Describe. | Grace:  will propose a tighter definition of Technical Variation and Measurement Variability (is this Standard Deviation) and clarify where reader variability fits in (or doesn’t) in this scheme.  <This Homework moved here from Comment 146>  David:  will provide clarifying language (in section IV) on exactly what it means to “pass” the +-15% variability requirement. (i.e. roughly what is the procedure. Is it “measure all the flagged tumors in a recognized test set, compute a volume change for each, add up all the absolute volume changes and see if you are within 15% of the value published with the dataset?” or “measure one tumor in one set twice and see if it’s within 15%” or “make up a definition of a some value you can demonstrate to be lower than 15 and claim compliance”  )  David’s Response: This is a question for the statisticians, but I would imagine that one would need something inserted at around Line 271 (which should be entitled “Image Analysis” rather than “Software Analysis Tool” to be consistent with other sections) that said something like “For the designated test image set of human subject images of n lesions, x repeat measurements by a single fully automated or y multiple human semi-automated or manual readers shall result in values within 30% (the claim) of each other z% (95% ?) of the time”. Alternatively, as suggested by TSB in comment #164, we might want the performance to be relative to a specified (expert consensus derived) “truth” value, but that is not really what the claim is (the claim is repeatability, not accuracy).  See Also Comments #127 & 128 & 164 |
|  | TSB | H | 258 | 4 | Lesion Volume Change:  Coming up with an effective performance metric for validation and comparison will be challenging.  Presumably a validation test will involve presenting the Analysis Tool with a collection of cases, measuring it’s volume estimate against an accepted “correct” result and aggregating the errors into an overall score.  Some experience has shown that some tumors might be described as having “linear” performance (i.e. slight changes in parameters or algorithm result in slight changes in accuracy) while others are “chaotic” (i.e. slight changes in parameters can result in large swings in accuracy or complete failure).  E.g. a narrowing of the tumor might be on a fine line between recognizing it as connected or distinct. So results for such touching tumors are bi-modal.  And of course different algorithms will be chaotically sensitive to different tumors.  Not that we shouldn’t try to come up with performance metrics, just that we should be prepared to tackle some interesting issues. | Performance scoring should take into account the “difficulty” of some tumors, either by weighting the errors from them lower, or by having a score for “the linear/easy ones” and a score for “chaotic/hard ones”. Otherwise the performance scores will tend to be chaotic as well and have lower predictive value.  Might classify test tumors by expert consensus or cross-product against the StdDev of the early scores on that tumor. | David:  <For validation QIBA might select a reference data set. That set will contain a variety of tumors (hard and easy). The profile text does not need to describe the test datasets in detail (but should indicate where test sets can be obtained?)>  David’s Response: “We will need a test set that a) is acquired according to this profile and b) contains designated lesions that span multiple body parts such that it is sufficiently representative; i.e., includes not just lung but also mediastinal, liver, adrenal, neck, axillary, mesenteric, retroperitioneal, pelvic, etc. masses of various different solid tumor types (lung, breast, colon). Such a public dataset does not currently exist as far as I know.” |
|  | TSB | H | 258 | 4 | Lesion Volume Change: It is not discussed/addressed directly, but it seems like many readers will assume (or maybe it’s just me) that the Analysis Tool at follow-up time will receive the volume result from the baseline scan, and will only segment the followup scan and take the difference. A source of variability is then that the algorithm/assumptions for the baseline segmentation may be different than those for the followup segmentation.  One paradigm for addressing this is for the Analysis Tool at follow-up time to always re-segment the baseline images (and perhaps any intermediate images) so that the algorithm bias and weaknesses are the same. On the downside, if the algorithm requires more than minimal intervention, it increases the workload on the operator, potentially to the point of being impractical. | Decide whether this is permitted, encouraged, required. | Added text requiring the Tool be prepared to analyze both baseline and current together, and highlighting the various issues involved with doing it separately that we don’t address, thus the requirement to do it together.  Observations:  - the groundwork we base our profile on have been done with the same app used for both the baseline and the followup measurement.  - our groundwork data indicates interoperator variation is modest. 3A may provide inter-algorithm variability information  - oncology doesn’t randomize because RECIST calls for following target lesions. (Although you could still multi-step it to allow “blinded” measurement).  - FDG PET group finds having a quantification document set that records the “provenance” of the data is useful for future analysis |
|  | SF | H | 531 | App D | Settings compatible with compliance | GE Discovery HD750 sct3 @ 120 manual mA? Is this for chest only? What slice thickness will the images be read at? Why not use auto mA with a min and max? If asking for recons of 1.25 by 1.0 they will be degraded. | Mike:  Good questions. When we put this together we might have been thinking of thin slices. However, the profile allows images up to 2.5 mm thick. As for using auto mA with a min and max (which is GE’s way of doing dose modulation), we are not going to give detailed specifics as to how to use dose modulation for each protocol. However we could allow this under some additional wording (proposed above) where we suggest that the techniques are adjusted for patient size (and just not be specific as to how that is done). |
|  | SF | H | 536 | App D | Settings compatible with compliance | The reconstruction parameters suggest a DFOV of 35 cm? What if the patient is large? | Mike:  <will propose method/rationale for handling large patients>  see item 149 “A protocol compliant with the QIBA profile will include some description as to how the protocol should be adjusted appropriately for patient size”  we could add “This includes adjustments both in the acquisition (such as increasing mAs for larger patients) and reconstruction (such as using a larger DFOV for larger patients)”  otherwise Kevin:  will add text that patients too large for 35cm DFOV are outside the profile |
|  |  |  |  |  | Didn’t get a lot of review by Techs (or Rads outside the inner circle for that matter). It’s possible that the profile text is not easy to understand; or the profile requirements are not practical. | <Consider circulating to Techs and Rads to review the result of resolution for “practicality” and “understandability”>  <Consider this as part of the Trial Implementation feedback, i.e. don’t delay publication for this.> |  |
|  |  |  |  |  | It’s not fully clear how one publication of the Profile differs from another and what, if any, the criteria are for publication. | Consider that the Profile gets published several times for several purposes. Need to pick good names for each case and clear description of the purpose for which it is intended and criteria for publication.  Public Comment: circulated for feedback, not for use  Trial Use?: Circulated for use, the committee has resolved issues that might prevent compliant actors from achieving the claim, but no “real world” data yet.  Veted?: Circulated for use, there is “real world” data collected using profile compliant actors that corroborates the claim | Kevin:  Drafted Strawman Proposal based on Strategic committee discussion. |

*Thank you for your comments!*