QIBA Lung Nodule Assessment in CT Screening Task Force
10 February 2015 at 23 PM CT {special time}
Notes provided by Dr. Gierada

In attendance:		**RSNA:**
David S. Gierada, MD (Co-Chair)	Rick Avila, MS	Joe Koudelik
Samuel G. Armato, III, PhD (Co-Chair)	Nicholas A. Petrick, PhD	Julie Lisiecki
James L. Mulshine, MD (Co-Chair)	David F. Yankelevitz, MD	

From: Gierada, David [mailto:gieradad@mir.wustl.edu]
Sent: Wednesday, February 11, 2015 12:56 PM
To: Julie Lisiecki; Samuel G. Armato; 'James L Mulshine'; Rick Avila; Petrick, Nicholas A; Andrew Buckler (Elucid Bioimaging) (andrew.buckler@elucidbio.com); dfyank@gmail.com; daniel.sullivan@duke.edu; efjackson@wisc.edu
Cc: Norma Sandoval; tatiana.johnson@duke.edu; jmkroneberg@wisc.edu; Joseph Koudelik; Fiona Miller

Subject: Small nodule limited group call yesterday, 2/10/2015

Thanks again to everyone for participating in the call. The main purpose was to discuss with Nick Petrick how the small and large nodule profile claims compare. Here's a summary (in black) of the discussion of each of Nick's previous email comments (in blue):

1. The large nodule claim has gone away from a symmetrical range (e.g., +_-30%). We are now using "There is a 95% probability that the measured change -25% to +30% encompasses the true tumor volume change." This accounts for the different fractional change for expanding vs. contracting lesions (or defining change relative to the mean size which isn't very clinically pleasing). You may want to consider using a similar approach.

   The small nodule claim range is symmetric because the confidence intervals were rounded by 1-3% for ease of use. In most cases they were rounded up and the actual confidence limits are slightly smaller than stated in the claim. Nick agreed that this approach is reasonable.

   It was noted that the large nodule confidence limits are based on reproducibility (change in scan and/or analysis conditions is allowed) so would be expected to decrease the more conditions are the same. The small nodule confidence limits are based on repeatability (scan and analysis conditions must be the same at both time points) and would be expected to increase as conditions change.

   The issue of bias also was brought up here and it was felt that it is OK to expect that bias in small nodule volume measurements is corrected for by the software analysis algorithm so is assumed to be zero. However this only holds for nodule shapes covered by the profile (must be spherical or nearly spherical). Bias of varying degrees would be more likely for nodules with more complex shapes, even with the bias corrections built into the software algorithm.

2. You'll also notice the large nodule claim is worded a bit differently to say that the true change falls within the interval. This places the claim in the context of the estimated difference. As I read the small nodule claims, it appears to be related to an individual volume measurement. Nothing wrong with this and if fact much of the phantom data is in the context of single volume estimates but it may not fit exactly with what is of clinical interest. You can talk a look at both formats and decide if one is more appropriate for the Profile.

   The large nodule claim wording read by Nick sounded appropriate and we will revise the wording of the small nodule claim.

3. As far as compatibility with the large nodule claim, it does not seem to meshes completely in the surface, especially for the 10-12 mm nodules. Since we generally think that larger nodules have less error, the large nodule claim sort many people might expect the 10-12 mm nodules to just meet the -25% to +30% limits with large lesions having smaller errors (as your claims do). The small nodule claim indicates these can be measured

within +-15%. This is quite a bit smaller than the large nodule claim.

I said this appears to conflict on the surface. For one thing, we know that larger complex nodules, say 50 mm or larger are actual quite difficult to measure reproducibly , in general, since then tend to be very complex and run into/through other structures. Therefore, the error in these large nodule measurements may be larger than for small, potentially more isolated 10-12 mm nodules. Also, the large nodule allows for complex shapes where the small nodule is limited to approximately spherical nodules. This would negative impact both bias and reproducibility leading to a larger range for the LN claim. Note, the complex shapes still need to be measurable in the LN claim (tumor margins are sufficiently conspicuous and geometrically simple enough to be recognized on all images in both scans; the tumor is unattached to other structures of equal density). In addition, the small nodule claim define measurable as being able "to be segmented using automated software without manual editing." This is a much more stringent condition than for the LN claim. Therefore, the two really aren't in conflict. You still may want to consider including something on compatibility between claims or modify the claims somewhat so that the a lay reader won't walk away confused.

We clarified that the small nodule claim for 10-12 mm nodules is that change can be measured to +-40% (+-15% was for the single volume measurement). This is still in conflict with the large nodule claim of -25% to +30%. Because there are multiple reasons that the numbers could be different, Nick and others felt the small nodule claim should be based on the small nodule data whether or not it is in perfect agreement with the large nodule claim.

4. I'd also suggest that you include an upper limit on claim 1d. You'll likely run into a similar problem with very large nodules (or more likely none will be able to fulfil your spherical and measurable criteria) as with the LN claim. They also likely aren't relevant to screening for the most part. I'm guessing something in the 20-30 mm range might be appropriate, but you'll need to decide if this makes sense or not.

   The group agreed with this recommendation, noting that as nodules get larger they are less likely to meet the profile specifications regarding shape and contact with structures anyway, and that in the screening setting the issue of growth becomes less important for larger nodules since they are usually biopsied.

   (After the call I noticed that the revisions made on last Friday's call I missed now limit the claims to nodules <12 mm.)

5. I also wonder if you limitation of spherical shape and measurability will severely limit the applicability of the Profile. You may want to include some examples of what are considered spherical enough even beyond long: short ratio since this is a very granular metric and likely won't be a sufficient metric. I'm also guessing that segmentation algorithms will still have problems with reasonable fraction of nodules, especially those attached/adjacent to the lung wall/vessels etc. Do you feel you have enough data to support that automated segmentation (I'm assuming the SN claim allows for manual seed placement/adjustment) is clinically viable. If you do, that would definitely be a nice step forward since we're still allowing for manual adjustment when necessary.

   The small nodule group feels that these restrictions are necessary in order to make the claims, recognizing that the profile won't be applicable to some nodules.

   Rick will try to come up with a limit for the nodule aspect ratio but the data may not be quickly extractable and he will be away in the coming weeks.

6. Now on to the numbers. Based on our phantom data results (we're just finished a "full analysis" of most of our phantom data with nodules ranging from 5-20 mm and using a matched filter volume estimator), I believe the SN claim numbers are consistent with what we found.

   A couple of factors to keep in mind which I'm sure you're already considering is the slice thickness, overlap

reconstruction and nodule complexity/density. With thicker slices and smaller nodules, overlapping reconstruction definitely help reduce the bias/variability in the volume estimates(See the attached papers for more details on both overlap recons and minimal detectable change for 5-10 mm nodules). As, I'm guessing you'll have more problems for low and/or mixed density nodules so some discussion of appropriate density may be important to consider as well.

Keep in mind, our results are only for phantom nodules so our variability/bias estimates are likely only a lower bound. Rick has more direct knowledge of what's achievable with small clinical nodules.

7. Do you have limits on the algorithm? For example are you going to require the same algorithm for all nodules. One concern I have is that different approach to volume estimation (e.g., voxel counting, model fitting) could lead to bias volume estimates of small nodules and reduced reproducibility. Consistency would suggest requiring the same algorithm. If multiple algorithms are used across time, you may have problems achieving the performance of the claim. I know that when intermixing 2 algorithms with the MSK data, the reproducibility increased from ~12% for the best algorithm to ~60% when mixed with any other algorithm. This was based on different data than yours so your results may/may not as much of an impact.

   It was felt that the performance expectations of the nodule analysis algorithm should be addressed in the compliance section of the profile. (I expect this will include a requirement that there is adequate bias correction.)

I reviewed the issues discussed on last Friday's call noted by Sam in his email and had a couple questions (highlighted).

3) Jim was going to prepare a statement regarding changes across size ranges -- in other words, how to handle the situation in Claim 2 in which a nodule is measured to be within one of the size ranges at time 1 and then is measured to be in another size range at time 2:

> I think we discussed that previously and concluded that the confidence limits to be used are those that apply to the size at time 1. Does the group now think that was the wrong approach?

4) Line 3 should include a statement specifying the metric that was used to obtain the "+/-" values in each of the subclaims (are these limits of agreement, or within-subject standard deviation, or something else?):

> I believe the numbers given by Rick are 95% limits of agreement (actually 95.45% since he noted in a previous email that he used 2 sigma to calculate the interval (2*sqrt(2)=2.83) rather than 1.96*sqrt(2)=2.77). To me this seems to be implied by the new wording of the claim, so is a separate statement needed?

We are clearly moving forward but at the risk of moving one step back, I found that it was hard to interpret the claims as worded in practical use. As an example: for a nodule in the 10 to <12 mm diameter range (524-<905 mm^3) at time 1, the claim says there is a 95% probability that the change measured at time 2 +/-40% encompasses the true volume change. Suppose one has a nodule that measures 10 mm (524 mm^3) at time 1 and at time 2 it measures 1024 mm^3 (nearly 100% increase in volume or nearly 3 mm increase in diameter). By my reading, the claim as written would tell the user there is a 95% probability that the true change is somewhere in the range of -700 to +700 mm^3. This range would encompass the possibility of a negative volume at time 1! I don't know if I'm interpreting the claim incorrectly but to me this doesn't seem very helpful. I know it's been said before but I think what one would like to know is how large of a %volume change is needed before it can be considered outside of the range of measurement variation with 95% confidence. Is there a way to give this guidance in the claim with wording that also has statistical validity?

The highlighted questions are for anyone with knowledge or an opinion-

Thanks,
David