Transcript: QMIC NIST talk

Daniel Sullivan

Slide 1: Title

Thank you, Dr. DiSalvo. I appreciate the invitation to tell you about the Quantitative Medical Imaging Coalition.

Slide 2: Disclosures

Slide 3: Why QMIC?

First, what are some of the issues that need attention? A couple of brief reminders from this morning that

[Advance]

clinical images are data sets,

So, we need to understand the sources of noise.

AI tools are involved, which brings into the methodology various issues related to AI image processing.

[Advance]

and although the concept of QI has matured over 25 years, clinical acceptance has been slow.

Slide 4: QMIC Work

QMIC‘s work is timely because:

[Advance]

measurements now drive care

[Advance]

standardization infrastructure is missing and

[Advance]

AI is only as good as the data

Slide 5: Founding

A summary about our founding. QMIC is a successor to the previous RSNA sponsored QIBA and QUIC activities, but with an expanded scope to cover the entire imaging chain. We were incorporated in March 2025 as a 501(c) non-profit,

[Advance]

QMIC was formed as a multi-stakeholder coalition from the start, not associated with any single professional society effort.

[Advance]

We think of our organization as a potential neutral convener, positioned as a trusted, neutral third-party to convene collective efforts across groups.

Slide 6: Mission

Our mission is to advance the science and implementation of quantitative imaging in clinical practice and clinical trials, and

[Advance]

Our goals are to develop QI guidelines based on rigorous metrology principles,

work towards clinical implementation of QI guidelines,

and collaborate with educational programs about QI.

Slide 7: Org Diagram

Organizationally QMIC is a coalition

[Advance]

with co-presidents Caroline Chung and Gudrun Zahlmann, both of whom had prior commitments today so could not be here in person,

[Advance]

a board of directors of nine people, of which I am one.

[Advance]

a steering committee of 30 made up mostly of the co-chairs of our biomarker committees.

[Advance]

and we hope to engage a broad range of related organizations and stakeholder groups to work together on relevant issues.

Slide 8: QMIC Standards

QMIC has 15 biomarker committees and importantly, we have a policy that each biomarker committee should formally collaborate from the start with a relevant clinical society so that the clinical perspective is intimately integrated into our Guidance Documents and not just provide an optional review and comment after draft documents are created.

[Advance]

Guidance Documents are international standardization papers that define how a quantitative imaging measurement should be made, reported and interpreted. More on that in just a minute.

Slide 9: Future plans

In terms of future aspirations,

[Advance]

we want to help develop conformance models,

[Advance]

centers of excellence or proficiency,

[Advance]

look for funding sources and

[Advance]

Participate in research collaborations.

Slide 10: what is a QMIC guidance document

Back to what is a QMIC guidance document? It describes a specific performance claim and how it can be achieved.

[Advance]

This approach is derived from the previous QIBA profiles and the structure is similar. The claim drives everything else in the document,

Slide 11: Claim

So, what is a claim? As noted already, it defines the quantitative performance of the guideline. It may or may not include a statement about bias, because you may remember from this morning, I said bias relates to the true value and in living humans it may not be possible to actually measure the true value.

But repeatability can be obtained under controlled conditions, in phantoms if necessary, and so a statement of precision defined as repeatability will always be a part of the claim.

Reproducibility may or may not be included because, like bias, it is sometimes difficult to get that in humans, for example if the imaging involves ionizing radiation. It's not ethical to expose the same person repeatedly to ionizing radiation just for the purpose of collecting repeat data under differing conditions.

Longitudinal change detection and the conditions required for those claims to hold are also included in a Claim.

Slide 12: Examples

Here are a couple of examples to illustrate how a claim is worded.

[Advance]

A cross-section claim is a single time point claim and

[Advance]

an example from PET would be

[Advance]

there is a 95% probability that the measured SUV max +/- 15% encompasses the true SUV Max

[Advance]

A longitudinal claim addresses how reliable a change measurement is over two scans

[Advance]

and example from CT tumor measurements would be

[Advance]

There's a 95% probability that the measured change in mass volume +/- 30% encompasses the true change in mass volume.

The ranges of +/- 15% and +/- 30% above are specific to the guidance documents that include these claims. They will be different for different clinical scenarios.

Slide 13: SLN Claims

To expand on this last point, here are claim statements extracted from the small lung nodule CT guidance document, to address a couple of questions that have come up about how precision values or coefficients of variation change depending on what size a lesion is or depending on other conditions under which the scan was obtained.

Claim number one is a single time-point measurement. It says "for a measured value Y and a coefficient of variation specified in table one, the 95% confidence interval for true nodule volume will be the result from this formula. We'll come to table one in just a moment.

[Advance]

Claim two is an example of a nodule volume change. You measure a nodule or mass at two time-points and calculate the percentage difference of the volume at time 2 compared to time 1 and call that X. The Claim then says, “The volume percentage change of X indicates that a true change in nodule volume has occurred if X is greater than the result of this formula, with 95% confidence.

Slide 14: Table 1

Here’s table 1 with various starting nodule diameters in column 1, the associated coefficients of variations in column 3 and the minimum detectable change listed in column 5. You can see that the coefficient of variation gets smaller as the nodule gets larger and the minimum detectable change gets smaller also.

If a given clinical setting needs a different minimum detectable change, for example, then the specifications in the Guidance Document would be adjusted to result in whatever measurement precision was needed or desired for that clinical application.

Also, as an aside, the precision required in a clinical trial would likely be different from the precision needed for individual patient management decisions. In a clinical trial it is necessary to detect a difference between two groups, and the assumption when designing the trial is likely to be that the difference between groups will be substantial. So less precision may be needed to detect this group-level difference, than would be needed to make an individual patient decision.

Slide 15: Conditions change

I was asked to address what happens when imaging conditions change across serial scans on the same patient. How does that affect precision? These examples come from the QIBA CT Volumetry Profile. At the top of this table is listed the possibility of different scanners or same scanner, then below that different or same radiologist, then different or same analysis

tools. In column 1 there are three examples of tumor size. We'll look just at the 50mm mass diameter line.

[Advance]

This would be when the Patient comes back and is scanned on the same scanner, same Radiologist, and same software. The minimum detectable change will be 8%. Very Good. We can be confident when measuring a very small change.

[Advance]

A more likely clinical setting is the patient comes back and is scanned on a different scanner but might get the same radiologist and probably the same software. Now the minimum detectable change is 24% so it's less good than the previous example, but not too bad – at least for cancer where size changes tend not to be trivial. But may not be very good for other diseases.

[Advance]

But very often, especially if the scans are done several months or a year apart, there would be a different scanner, different radiologist, and different software. Now the minimum detectable change is 43% so it goes up considerably in this example. The actual magnitudes of these effects would differ depending on the clinical scenarios for which a given Guidance Document is created.

Slide 16: Fleischner

How would these guidelines actually affect clinical practice?

[Advance]

This is a table of management recommendations for CT pulmonary nodules, and you can see that for nodules less than 6 mm nothing has to be done, and if it's greater than eight there's some intervention that needs to be done.

[Advance]

so, we could use 7 mm as a pragmatic threshold for a clinical decision of “wait” versus “intervene”

Slide 17: Overall Goal

This graphic shows an example of this common diagnostic decision scenario. There is a test result cutpoint or threshold that triggers a decision between, for example, treat versus wait. In this screen-detected nodule example, I’ve placed the threshold at 7.

[Advance]

If the measurement of a nodule is seven

[Advance]

but if the variability around that number is +/- 3, then we don’t know with high confidence whether the nodule is actually less than six and we should wait, or more than eight and we should treat. Adherence to the technical guidelines influences this decision-event in this way.

[Advance]

First there has to have been an analysis of all the sources of variance, and Kishore Rajendran is going to say more about this topic shortly.

[Advance]

The guidelines will have specified all the parameters that need to be controlled to minimize that variance down to the level that’s needed for this clinical setting, which in this example would have to be +/- 1.

[Advance]

When all of that is in conformance, and we do the measurement

[Advance]

then we will be within a range of precision where

[Advance]

we can have confidence that our clinical decision is consistent with a reliable measurement.

[Advance]

That is, we are making the correct clinical management decision.

Slide 18: Challenges 1.

Switching topics now to: What are some challenges to widespread implementation of rigorous quantitative imaging biomarkers? An important one could be called “inertia”, or embedded culture. Clinicians and radiologists think of Imaging as a picture not an assay.

[Advance]

There’s an expression called the Lindy effect: that “The longer an idea has been around the longer it tends to persist. Clinical images have been thought of as pictures for over 100 years. This deeply ingrained idea is very difficult to change.

[Advance]

Thus we need robust educational advocacy efforts, especially from professional organizations but, frankly, they don’t have major incentives to do this.

Slide 19: Challenges 2

There are several additional impediments to implementation, and here are four that I think are important.

[Advance]

One is that both users and payers say, “show me the data that there’s a benefit”, so we need outcome data associated with the QIBs, and these are difficult to obtain.

[Advance]

We need industry agreement on cross-vendor harmonization and interoperability. There are examples that this is possible. Not surprisingly, there’s some hesitancy among manufacturers for various reasons, but it is possible to get past this.

[Advance]

To overcome this range of friction points we need triggers to prompt clinicians and users to adopt QIBs. They could be political, reimbursement, regulatory or other incentives and I'll give you two examples.

[Advance]

Overall, for everyone to have confidence in the clinical value of these measurements, we need QA evaluation and accreditation processes, which currently do not exist.

Slide 20: Political Trigger

The European CT lung cancer screening programs are examples of political and regulatory triggers.

[Advance]

Lung nodule management decisions are driven directly by measurements as I showed a few moments ago. CT screening QIB standards began as a QIBA Profile and has been revised as a QMIC Guideline.

[Advance]

In 2023,the EU funded a program known as SOLACE, to develop programs throughout Europe to promote CT screening.

[Advance]

SOLACE in turn created the European lung cancer screening alliance, which is managed jointly by a radiology and a clinical organization.

Slide 21: ELCSA

The European lung cancer screening alliance exists to promote lung cancer screening across all EU countries. An important activity is to harmonize standards, especially Quality Assurance.

Slide 22: German

The European activities I've described so far could be considered political. In addition, in some European countries, particularly Germany, it's also becoming regulatory.

[Advance]

There are now two laws in Germany covering CT screening. I'm not going to attempt to read these in German but one of them was in 2024 a more recent one, 2026.

[Advance]

Together these mandate specific QA requirements for lung cancer screening and radiologist qualifications, QA requirements for the CT equipment, for the reading workstation software and interestingly, mandatory double reading and mandatory AI assisted nodule detection software.

Slide 23: German laws

These are examples from our QMIC involvement in Germany. One of our co-presidents, Gudrun Zahlmann, lives in Germany and provided these documents to me. They are all in German, but she has helpfully labeled them to indicate

[Advance]

that this one is for volumetric assessment that follows the QMIC guidelines,

[Advance]

this one for scan parameters comparable to the QMIC guidelines,

[Advance]

and this document with some technical parameters which don't yet conform to QMIC guideline. She's indicated this as a task for QMIC to work on.

Slide 24: MASLD

MASLD is an example where reimbursement was the important trigger.

[Advance]

I won't repeat what I talked about this morning but point out the third bullet. MR experts contend that the Blue Cross established eligibility criteria (that is the threshold or cut points chosen by the Blues) are not optimum for several reasons, and Modeling could very likely propose better cutpoints that would improve patient benefit, cost-effectiveness, and appropriate QA requirements.

Slide 25: Claim Precision

Lastly, this brings up a related topic which is, when writing a claim, what level of precision is needed?

[Advance]

Well, for any given setting it depends. . .

[Advance]

What is the current performance of the QIB? What's the range of plausible performance values? How good does the measurement need to be to inform a clinical decision? What's the minimum acceptable performance threshold.

[Advance]

Then, after the biomarker committee decides on a potential claim, What will be the sensitivity/specificity implications of a selected threshold? What are the clinical consequences of the false negatives and false positives that inevitably will result from whatever cutpoint is being considered?

Slide 26: Recommendation 1

I will conclude with two specific recommendations for the government agencies involved here today.

[Advance]

One is that we need a national QI metrology center of excellence for clinical imaging.

[Advance]

There is a detailed proposal that is available on the web, and

[Advance]

there are a couple of potential models from the laboratory world. The CDC has the lipid standardization program, which is relevant, and CMS has the well-known clinical laboratory improvement amendment program CLIA.

Slide 27: Recommendation 2

Secondly, NIST's foundational mission is measurement science — establishing traceable, SI-linked reference standards. They have deep expertise in statistical and algorithmic rigor, uncertainty quantification, and error propagation.

[Advance]

They are the most authoritative source for reference standards, including both guidance documents, and physical reference materials and phantoms.

[Advance]

And NIST is ideally suited to develop and lead scanner calibration reference facilities.

Slide 28: Summary

So, in summary I would just remind you,

[Advance]

that quantitative Imaging is an assay.

[Advance]

QI must be grounded in fundamental metrology principles. That’s a basic QMIC premise.

[Advance]

Professional organizations must educate users to implement QI as an assay. QMIC will assist with this.

[Advance]

and federal agencies can provide standards and QA infrastructure.

Slide 29: Thank you

Sources & References:

www.QMIC.org

info@qmic.org

linkedin.com/company/quantitative-medical-imaging-coalition

Federation of American Scientists (FAS) Publication (2024)

“A Quantitative Imaging Infrastructure to Revolutionize AI-Enabled Precision Medicine”

fas.org/publication/ai-enabled-precision-medicine/

Two legal instruments covering CT lung screening in Germany:

- the Lungenkrebs-Früherkennungs-Verordnung (a federal ordinance from the environment/radiation-protection ministry, BMUV, July 2024)
- the G-BA's Krebsfrüherkennungs-Richtlinie (the Federal Joint Committee's directive, amended to cover lung screening, March 12, 2026).
-

SOLACE https://europeanlung.org/solace/

ELCSA https://europeanlung.org/solace/2026/02/03/solace-launches-the-european-lung-cancer-screening-alliance-elcsa/