Quantitative Imaging Biomarker Alliance

APPROACH TO COMPLIANCE

August 15, 2011

Andrew J. Buckler, MS

Program Director, QIBA

PRINCIPAL LOGISTICAL AND FINANCIAL SUPPORT PROVIDED BY RSNA

An Important Aspect of the QIBA Value Proposition…

- Efficiently collect and exploit a body of evidence establishing standards for optimized QI:
  - Users want confidence in the read-outs
  - Pharma wants to use them as endpoints
  - Device/SW companies want to market products that produce them without huge costs
  - Public wants to trust the decisions that come from them

Converging a methodology to assess compliance helps with realizing this value for our stakeholders.

Descriptive Statistics

- Technical Performance (based on such studies as 1a-c, 3A)
  - Bias
  - Variability
  - Reliability
    - Repeatability: repeat measurement under same conditions
    - Reproducibility: repeat measurement under different conditions
- Clinical Performance (based on such studies as 3B and RCTs)
  - Validity
    - Predictive value (positive/negative)
    - ROC (sensitivity/specificity)
  - Value as a Surrogate
    - Effect of treatment on true endpoint, treatment on surrogate endpoint, surrogate on true endpoint
    - Proportion of treatment effect on true endpoint explained by surrogate

Profile details give formal definition for activities

User Perspective

Will it do what I need?

What/who do I need to get started?

What do I have to do (procedures, training, performance targets) to achieve the Claims?

Claims:

"Detect tumor response with twice the sensitivity of RECIST in the Lung"

Details:

Actors Table

- CT Acquisition System
- Measurement Software
- Radiologist

Activity Definitions

- Calibration / QA
- Patient Preparation
- Image Acquisition
- Reconstruction
- Post-Processing
- Analysis / Measurement
- Reading / Interpretation
- …

Vendor View

Why do you want me to do this?

Which of my products are affected?

What do I have to implement; (features, capabilities, performance targets)

How will I be tested?

Figures of merit differ for each activity
(relevant to all levels except “none”)

	Figures of merit	Experimental results (descriptive statistics)
Acquisition	Noise, contrast, reconstructed pixels	Regional histograms and specific reconstructed pixel values
Post-processing	Determinacy, dynamic range	Non-overlap in range over the defined domain
Analysis	Accuracy, precision, reproducibility	Inter- and intra-reader bias and variance
Interpretation	Sensitivity and specificity	ROC analysis, hazard ratios, K-M curves
Storage	Enumerated fields	Check for the fields

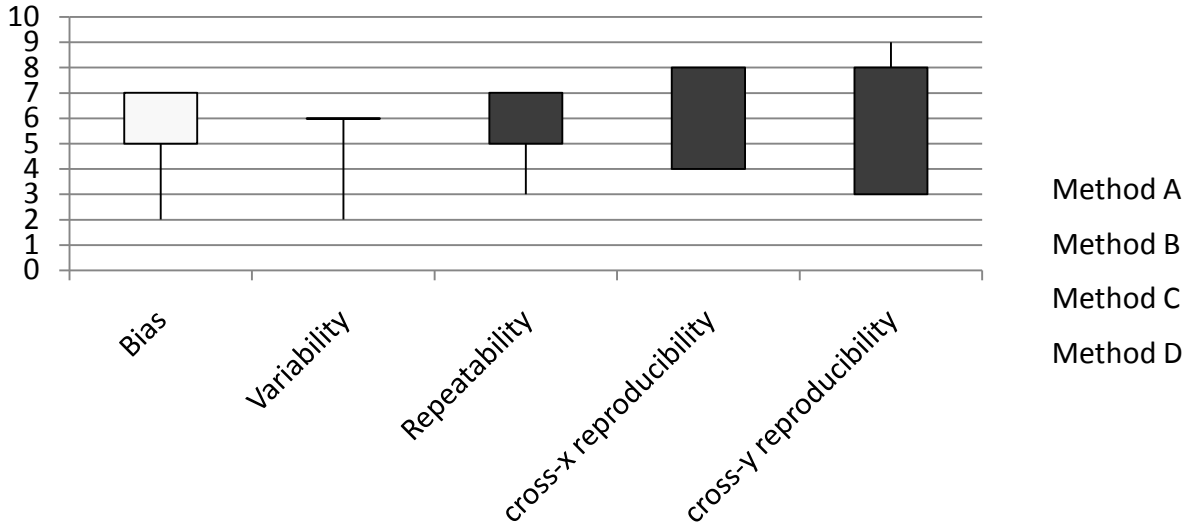
Compliance approach for each activity
(relevant to all levels except “none”)

	Figures of merit	Experimental results (descriptive statistics)	Compliance approach
Acquisition	Noise, contrast, reconstructed pixels	Regional histograms and specific reconstructed pixel values	Physical-standards-traceable phantoms with serial numbers, supported with shipping transactions where shelf life is an issue, where a scripted acquisition and data transmission is performed
Post-processing	Determinacy, dynamic range	Non-overlap in range over the defined domain	Similar to acquisition except that a DRO is used in place of a physical phantom, and similar to Analysis in terms of ability to utilize a black-box wrapper
Analysis	Accuracy, precision, reproducibility	Inter- and intra-reader bias and variance	Black-box algorithm wrapper with optional human reader steps in batch analysis as a web service over a sequestered reference data set specified by context for use claims
Interpretation	Sensitivity and specificity	ROC analysis, hazard ratios, K-M curves	As with analysis (though in practice weighting shifts from algorithm to reader)
Storage	Enumerated fields	Check for the fields	Incorporated within above

Step 1: perform initial set of methods on reference objects, assessing performance of each one with respect to the selected descriptive statistics

	Bias	Variability	Repeatability	cross-x reproducibility	cross-y reproducibility
Method A	5	6	7	8	8
Method B	6	2	3	4	9
Method C	2	5	6	7	5
Method D	7	6	5	4	3

The dispersion in performance of these methods with respect to the selected descriptive statistics may be visualized using box plots.

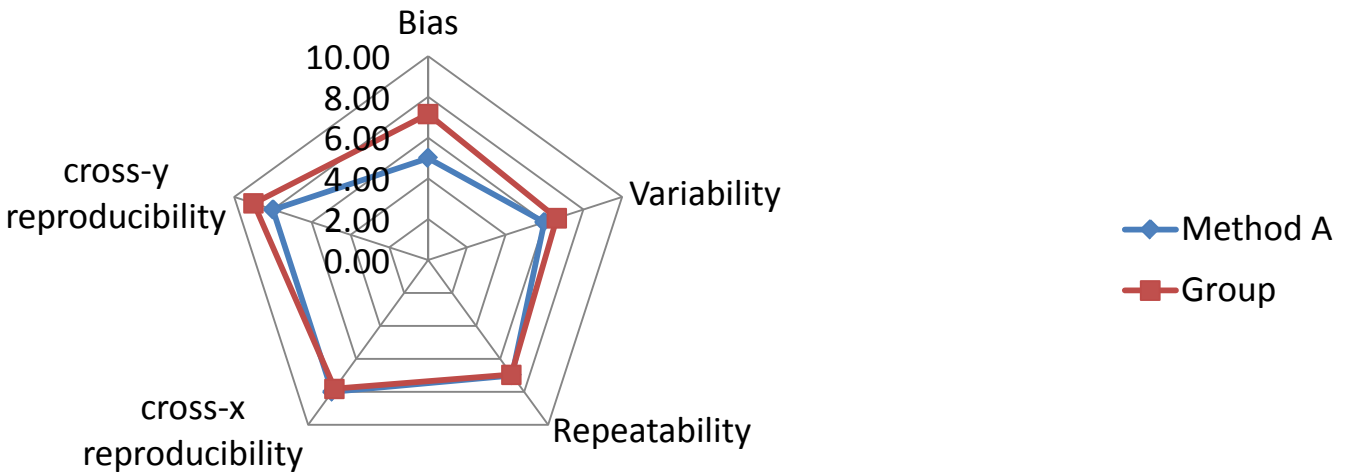


Step 2: Select a group value for each of the descriptive statistics, e.g., as the mean plus 1 stdev (or as wide as we think wise)

	Bias	Variability	Repeatability	cross-x reproducibility	cross-y reproducibility
Average	5.00	4.75	5.25	5.75	6.25
Stdev	2.16	1.89	1.71	2.06	2.75
Group value	7.16	6.64	6.96	7.81	9.00

Step 3: Compose a radar plot for the group value and plot any given method with respect to it.

	Bias	Variability	Repeatability	cross-x reproducibility	cross-y reproducibility
Method A	5.00	6.00	7.00	8.00	8.00
Group	7.16	6.64	6.96	7.81	9.00

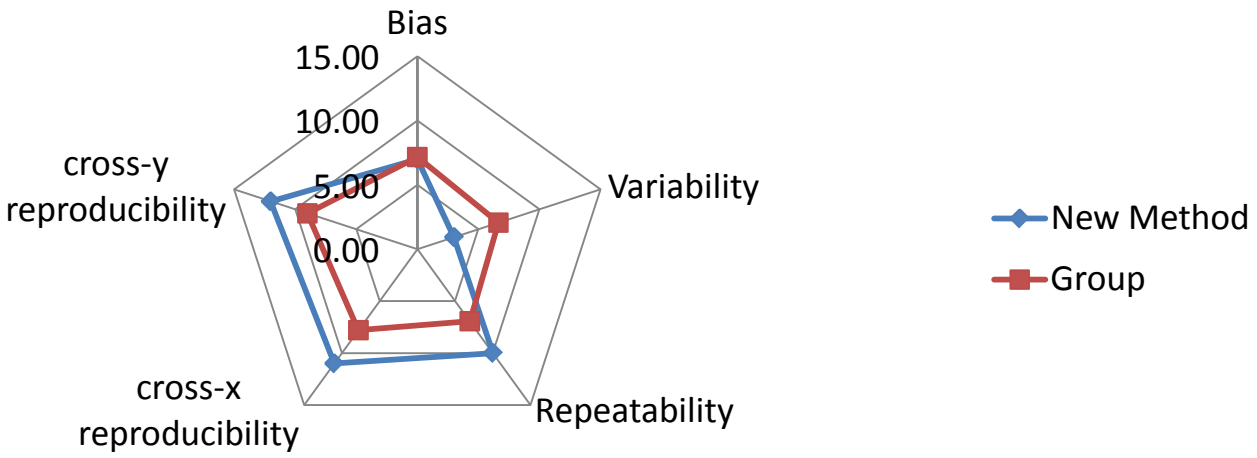


In this example, Method A could be said to be “compliant” in that its performance fits within the group values.

Step 4: As new methods come along that need to be assessed, they can be run on the same reference data set and their performance may be compared to the group values.

In this example, the new proposed method does not perform well enough to be considered a valid method since it falls outside the group values.

	Bias	Variability	Repeatability	cross-x reproducibility	cross-y reproducibility
New Method	7.00	3.00	10.00	11.00	12.00
Group	7.16	6.64	6.96	7.81	9.00



One way to think of compliance is to pick among a level that we shoot for

- None
- Self-test
- Self-certify
- Certificate obtained from 3rd party (e.g., QIBA)

Each level has different strengths, and each requires a different infrastructure.

3rd party certification is mostly the same as the other levels but adds:

- Formal sequestered data through trusted broker
- A certifying organization into the workflows (e.g., probably QIBA)
  - Optionally can sub-contract proficiency testing services from NIST on a cost-recovery basis
- Can facilitate formatting of data to be seamless with agency in setting up a “master file” on behalf of sponsors who wish to reference it

Let's talk about our approach to compliance

- Value proposition of our effort includes efficient collection and exploitation of a body of evidence establishing standards for optimized QI
- Compliance plays a role in this
- Our Profile Claims provide a base from which to work in a disciplined fashion:
  - it specifies the read-outs,
  - it articulates testable hypotheses,
  - it identifies the biology,
  - it sets out the purpose, and
  - it references the defined method that is elaborated in the Profile details
- The Profile Details give formal definition for the activities, allowing figures of merit to be established in a way that can be measured
- Support for establishing tools and methods are being developed with sin the QI-Bench program