

# Report of Ground Truth Task Force

**Task Force Members:**

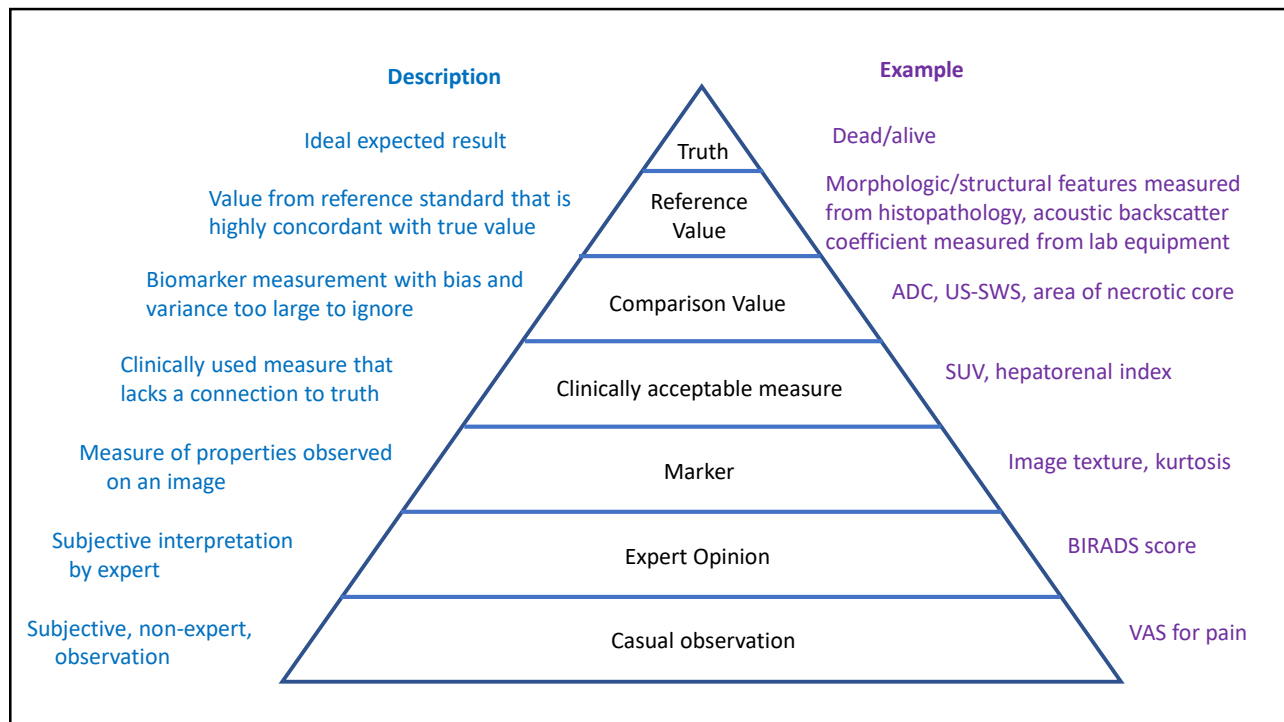
Andrew Buckler, MS  
Jana Delfino, PhD  
Alexander Guimaraes, MD, PhD  
Tim Hall, PhD  
P. David Mozley, MD  
Nancy Obuchowski, PhD (Chair)  
Nicholas Petrick, PhD  
Gudrun Zahlmann, PhD

1

## Task: Assess the role of “ground truth” in QIBA’s Profiles

- Following Sullivan et al [Radiology, 2015], the Task Force chose not to use the term “ground truth” but rather discussed “true value”.
- Task Force identified need to define a hierarchy of terms that included “true value”, “reference value”, “biomarker measurements”, etc. to delineate the relationships between these terms.

2



3

## Observations:

- True Value, although not always easily observed, is assumed to exist.
- Reference Value and Comparison Value are each defined relative to a true value. The pyramid levels below these may not have a relationship to a true value.

4

## Observations (continued):

- Biomarkers fall under Comparison Values where there is a clear connection to truth. Practically, though, a true or reference value may not be available for all biomarkers. When available, measurement bias can be estimated.
- When comparing a biomarker to another Comparison Value, agreement can be assessed, but the bias of the biomarker cannot be estimated.

5

## Considerations:

- The ability to characterize a measurement relative to a true or reference value is important to quantitative imaging. This is possible with biomarkers, but not 'Clinically Acceptable' measures, nor imaging 'Markers'.
- QIBA has focused on biomarkers in the past and has established a respected and trusted methodology for writing Profiles for biomarkers.
- In contrast, 'Markers' lack a connection to truth; thus, performance metrics currently used by QIBA, such as bias and linearity, as well as quantification of true change over time, are not applicable to 'Markers'.

6

## Recommendations (1):

- The Task Force does not feel that imaging 'Markers' should categorically be excluded from QIBA; however, given limited resources and the fundamental differences in performance metrics, the Task Force recommends that QIBA prioritize Profiles involving biomarkers over Profiles involving markers.

7

## Recommendations (2):

- In the past QIBA has included biomarkers where no reference value is available, focusing on longitudinal claims in the Profiles. The Task Force does not see a need to change this approach.

8

## Recommendations (3):

- Multiparameter quantitative imaging (mp-QI) is of great interest to the clinical community, as well as to QIBA.
- The Task Force recommends that QIBA prioritize biomarkers in their mp-QI Profiles, over markers.
- Furthermore, since calibration is an important metric for describing the performance of prediction models, QIBA should prioritize models whose outcome has a relationship with a true value so that calibration of the model can be reported.