An Overview of Current Progress towards an Evidentiary Framework for Biomarker Qualification

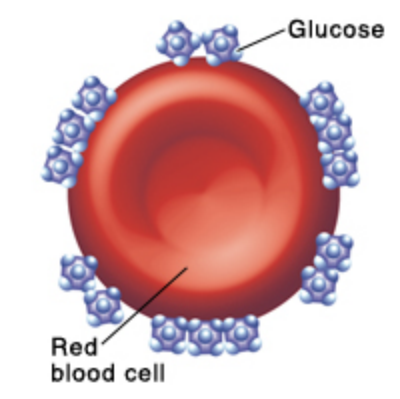
John A. Wagner, MD, PhD
Takeda Pharmaceuticals International Co.
Clinical and Translational Science
The Biomarkers Consortium

Outline

- Biomarkers and the tower of Babel
- 48 BEST definitions
- 1 recent example
- 5 step qualification framework
- BEST + 5 step = faster, more efficient qualification?

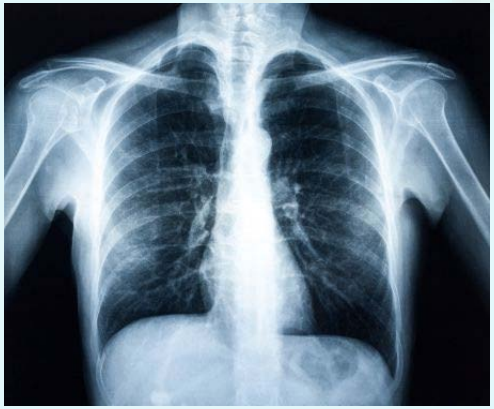
Biomarkers

We use biomarkers all the time in clinical practice and drug development

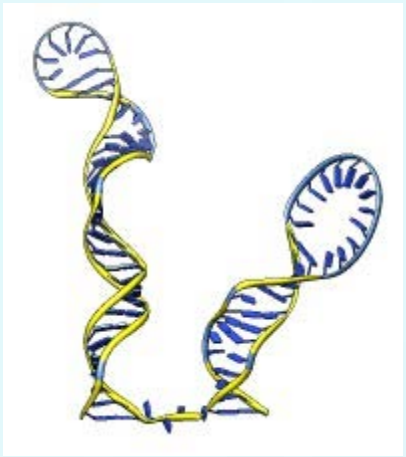


Hemoglobin A1C

Blood pressure



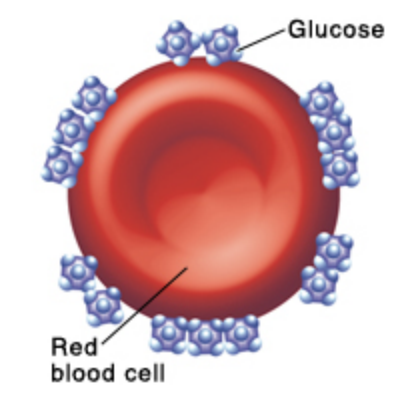
Radiographic evidence of tumor shrinkage



HIV-RNA reduction

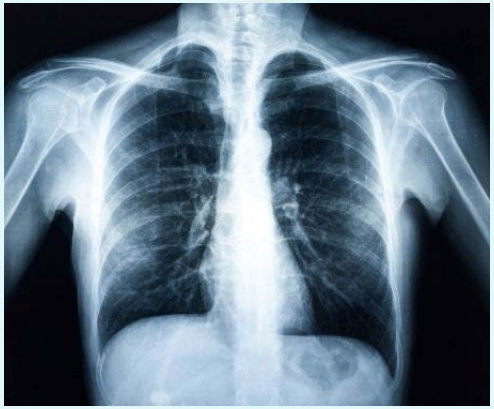
Biomarkers

These are all validated or reasonably likely surrogate endpoints…

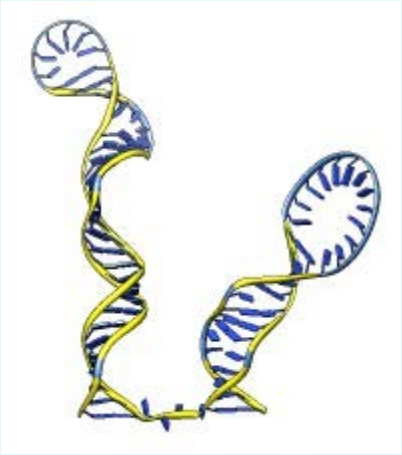


Hemoglobin A1C

Blood pressure

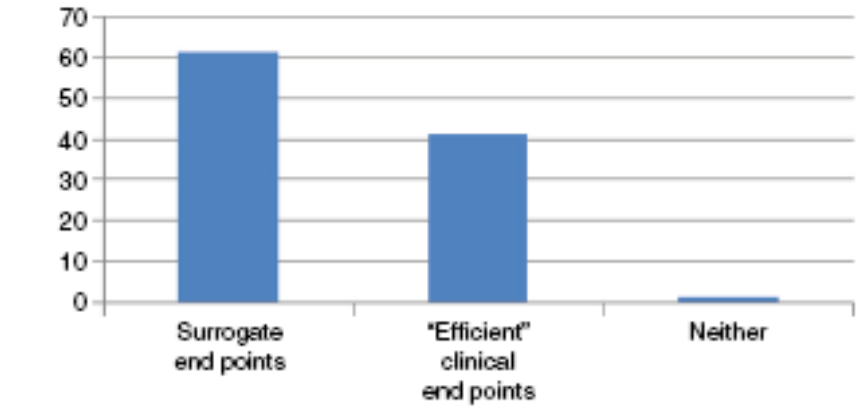


Radiographic evidence of tumor shrinkage



HIV-RNA reduction

Biomarkers are critical in drug development



- Hypertension (240)
- Diabetes (112)
- Cancer (>100)
- HIV (61)
- MS (63)
- Cholesterol (50)
- Asthma (30)

- Rheumatiod arthritis (>100)
- Rhinitis (>200)
- Back pain (>100)
- Parkinson's symptoms (98)
- Depression (60)
- Schizophrenia (47)
- Anxiety (37)
- ADHD (28)
- Obesity (24)
- Urinary incontinence (15)
- Alzheimer's symptoms (12)
- Myasthenia gravis (11)

- Age-related macular degeneration (6)
- ALS (2)
- Alzheimer's neurodegeneration (0)
- Stroke neuroprotection (0)
- COPD progression (0)
- Huntington's (0)
- Liver fibrosis (0)
- Diabetic complications (0)
- Osteoarthritis progression (0)

- Surrogate endpoint use in drug approvals
  - Simple survey with WebMD
  - Surrogate endpoints associated with higher numbers of new drugs when compared with similar conditions for which they do not exist
  - "Efficient" clinical endpoints similar to surrogate endpoints

Lathia et al. CPT, 86:32-43, 2009 PMID:19474783

The biomarker tower of Babel

Language confusion hinders medical practice and drug development

- Misinterpretation of evidence
- Misunderstanding of evidentiary requirements
- Failure of clinical trials
- Delays
- Potential harm to patients

- What is the difference between a surrogate endpoint and surrogate marker?

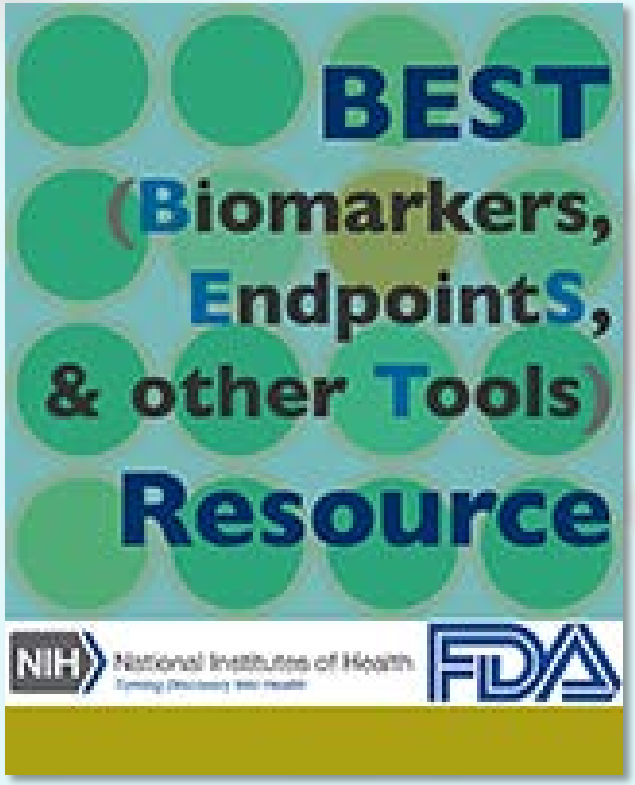
The biomarker tower of Babel

Language confusion hinders medical practice and drug development

- Misinterpretation of evidence
- Misunderstanding of evidentiary requirements
- Failure of clinical trials
- Delays
- Potential harm to patients

- What is the difference between a surrogate endpoint and surrogate marker?

BEST: BIOMARKERS, ENDPOINTS, AND OTHER TOOLS RESOURCE



- A glossary of terminology and uses of biomarkers and endpoints in basic biomedical research, medical product development, and clinical care
- Created by the NIH-FDA Biomarker Working Group
- Publicly available at http://www.ncbi.nlm.nih.gov/books/NBK326791/
- Periodic updates planned with additional terms, definitions, and examples.
- Feedback welcome (biomarkers@ncbi.nlm.nih.gov)
- Published January 28, 2016
- Last Update: December 22, 2016.

BIOMARKERS

Definition: A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.*

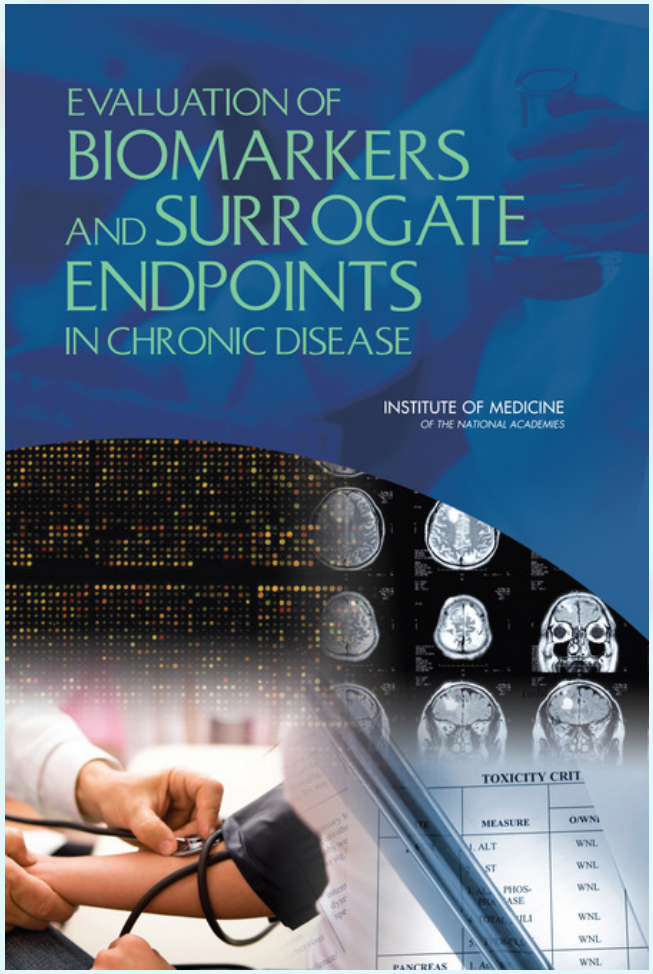
Types: Molecular, histologic, radiographic, or physiologic characteristics are types of biomarkers

Example: pharmacodynamic/response biomarker

Used to show that a biological response has occurred in an individual who has been exposed to a medical product or an environmental agent.

***Updated definition from* BEST Glossary**:
http://www.ncbi.nlm.nih.gov/books/NBK326791/

Surrogate endpoint



- An endpoint that is used in clinical trials as a substitute for a direct measure of how a patient feels, functions, or survives. A surrogate endpoint does not measure the clinical benefit of primary interest in and of itself, but rather is expected to predict that clinical benefit o harm based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence
- From a U.S. regulatory standpoint, surrogate endpoints and potential surrogate endpoints can be characterize by the level of clinical validation:
  - validated surrogate endpoint
  - reasonably likely surrogate endpoint
  - candidate surrogate endpoint

Qualification vs. validation

Analytical validation: Establishing that the performance characteristics (including sensitivity, specificity, accuracy, and precision) of a test, tool, or instrument are acceptable.

Clinical validation: Establishing that the test, tool, or instrument acceptably identifies, measures, or predicts the concept of interest.

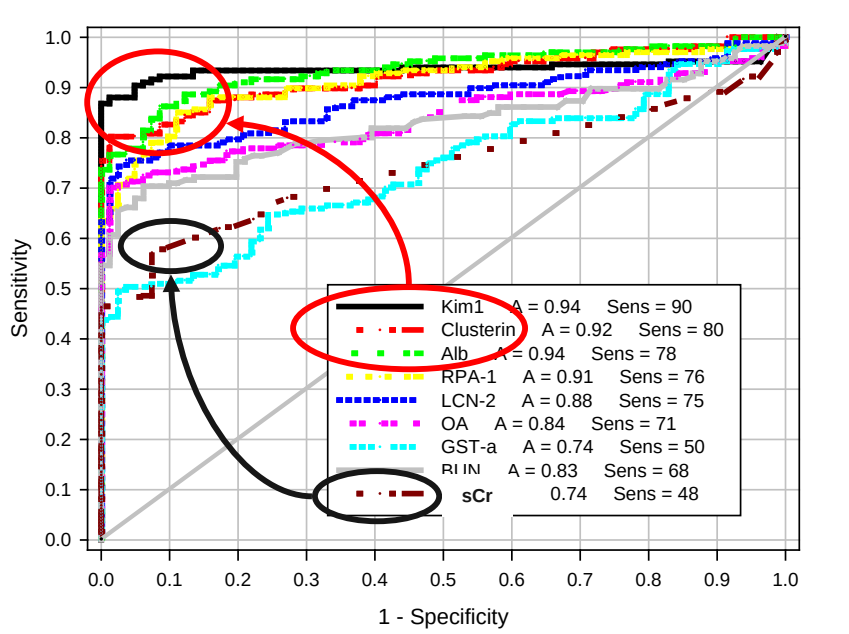
- Concept: In a regulatory context, the concept is the aspect of an individual's clinical, biological, physical, or functional state, or experience that the assessment is intended to capture (or reflect).

Qualification: A conclusion, based on a formal regulatory process, that within the stated context of use, a medical product development tool can be relied upon to have a specific interpretation and application in medical product development and regulatory review.

BEST Glossary: ***http://www.ncbi.nlm.nih.gov/books/NBK326791/***

The promise and pitfalls of novel, translational kidney biomarkers

- New kidney safety biomarkers outperform serum creatinine and BUN in rats
  - FDA, EMA, PMDA Qualification



The Hypothesis: **New promising translational kidney safety biomarkers could:**

1) Mechanistic insight,
2) Earlier and more sensitively than BUN and sCr
3) report dysfunction AND damage
4) Inform patient prognosis
5) Enable safe clinical drug development

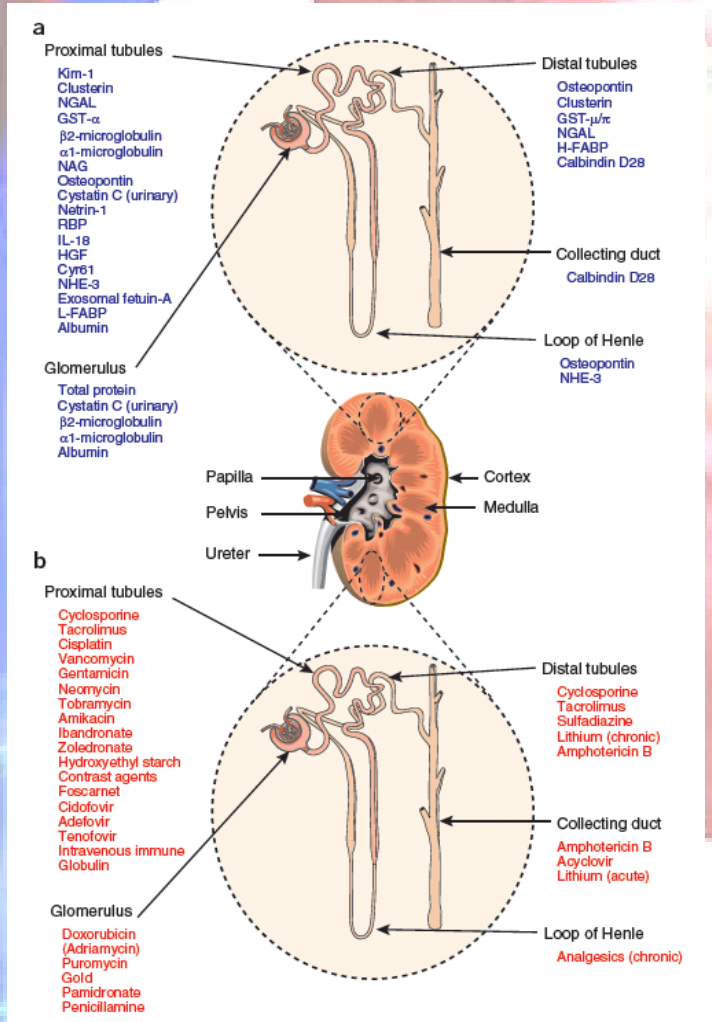


Figure 1 The utility of biomarkers to detect injury to specific nephron segments affected by various nephrotoxicants. (a) Nephron segment-specific biomarkers of kidney injury. (b) Drugs that elicit site-specific toxicity in the kidney[12,13].

Promising urinary biomarkers of acute renal tubular damage or dysfunction to complement BUN and serum creatinine

Functional Biomarkers	Proposed Functional Interpretations
Albumin	Small quantities filtered by glomerulus and efficiently reabsorbed by tubular epithelium
Cystatin C	Normally highly filtered but either glomerular or tubular damage yields protein overload that inhibits tubular reabsorption from lumen
Total Urinary Protein	Functional marker of glomerular filter integrity or tubular dysfunction
Injury Response Markers	**Proposed Structural Interpretations**
Clusterin	Necrotic tissue sequestration; and regenerative repair response present in many renal cell types
Kim-1	Tubular epithelium dedifferentiation and regenerative repair response
NGAL (Lipocalin 2)	Also filtered and reabsorbed; distal tubule inflammation and to sequester iron, limit damage.
Osteopontin	Expressed in TAL and DCT, may limit oxidative stress and ischemia, and assist regeneration
Leakage Markers	
NAG	Brush-border enzyme released when damage occurs to tubular epithelium

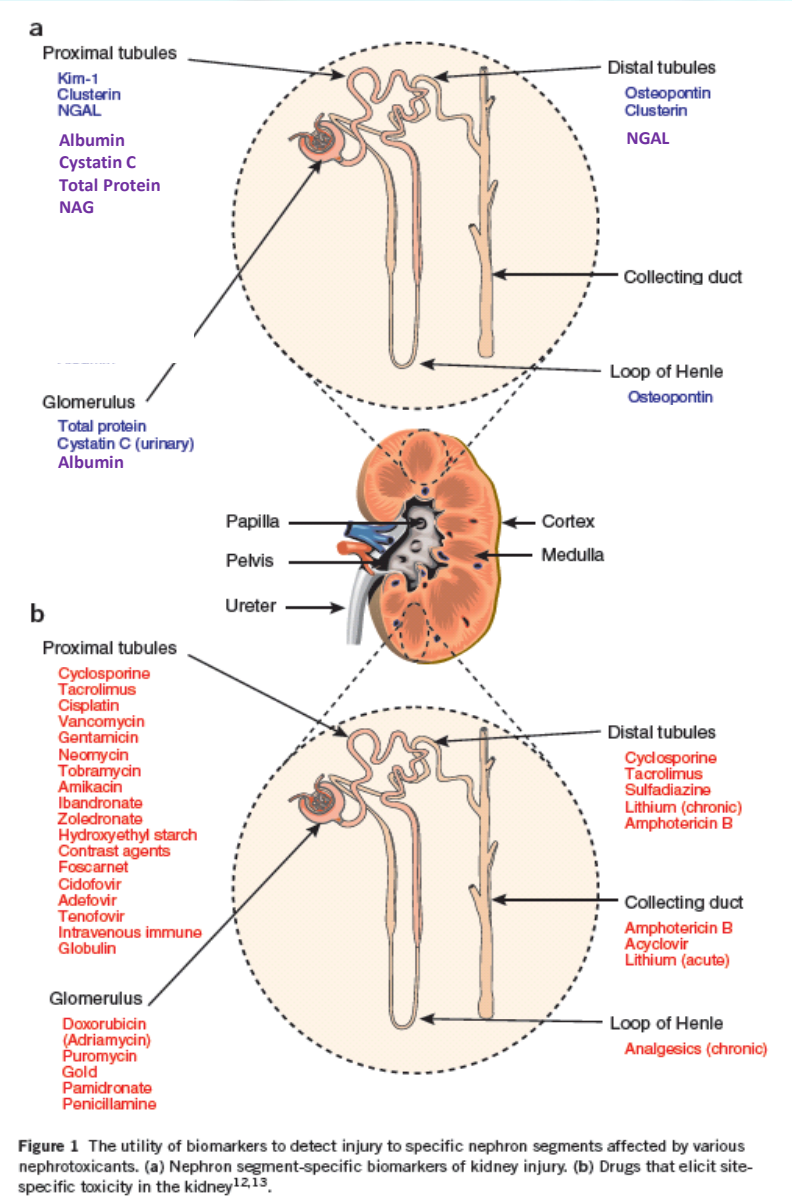


Figure 1 The utility of biomarkers to detect injury to specific nephron segments affected by various nephrotoxicants. (a) Nephron segment-specific biomarkers of kidney injury. (b) Drugs that elicit site-specific toxicity in the kidney[12,13].

Example: Published evidence supporting enhanced sensitivity of KIM-1 over sCr

AJKD
Case Report
Am J Kidney Dis. 2013;62(4):796-800

Acute Kidney Injury During Therapy With an Antisense Oligonucleotide Directed Against PCSK9

Eveline P. van Poelgeest, MD,[1] Reinout M. Swart, MD,[2] Michiel G.H. Betjes, MD,[2] Matthijs Moerland, PhD,[1] Jan J. Weening, MD,[3] Yann Tessier, DVM,[4] Michael R. Hodges, MD,[4] Arthur A. Levin, PhD,[4] and Jacobus Burggraaf, MD, PhD[1]

Antisense oligonucleotides have been explored widely in clinical trials and generally are considered to be nontoxic for the kidney, even at high concentrations. We report a case of toxic acute tubular injury in a healthy 56-year-old female volunteer after a pharmacologically active dose of a locked nucleic acid antisense oligonucleotide was administered. The patient received 3 weekly subcutaneous doses of experimental drug SPC5001, an antisense oligonucleotide directed against PCSK9 (proprotein convertase subtilisin/kexin type 9) that is under investigation as an agent to reduce low-density lipoprotein cholesterol levels. Five days after the last dose, the patient's serum creatinine level increased from 0.81 mg/dL at baseline (corresponding to an estimated glomerular filtration rate [eGFR] of 78 mL/min/1.73 m^2) to 2.67 mg/dL (eGFR, 20 mL/min/1.73 m^2),

A post hoc analysis of biobanked spot urine samples, which had been collected before each dose of study medication was administered, was performed to assess the kidney injury markers β_2-microglobulin, α-glutathione *S*-transferase (α-GST), kidney injury molecule 1 (KIM-1), and *N*-acetyl-β-D-glucosaminidase (NAG). NAG levels were unchanged, but urinary β_2-microglobulin levels increased 4-fold, α-GST levels increased 24-fold, and KIM-1 levels increased 60-fold upon administration of SPC5001 (Fig 2). Importantly, these markers preceded the increase in serum creatinine level, having increased already after the first administration of SPC5001. These observations suggest that SPC5001 adversely affects proximal tubular function.[15,16]

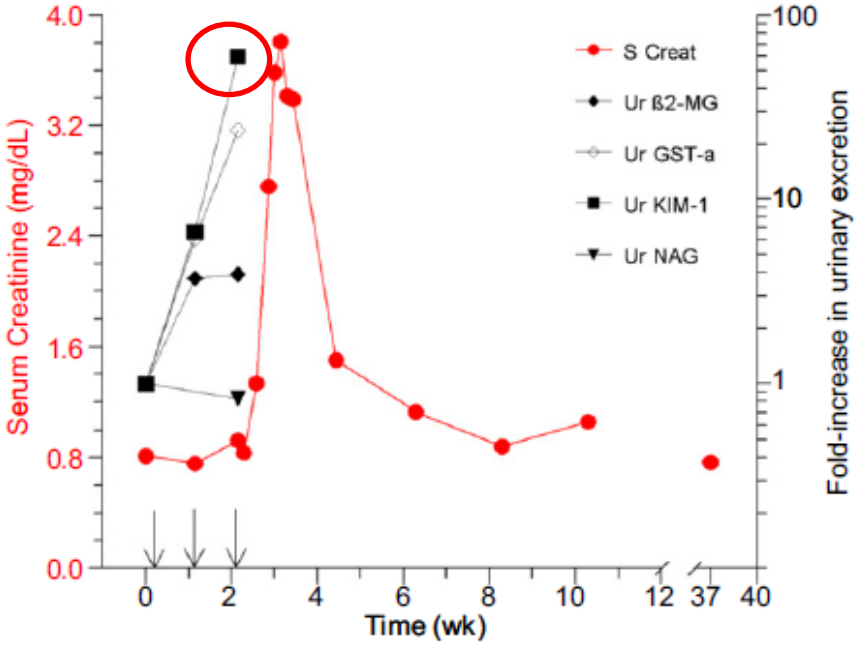
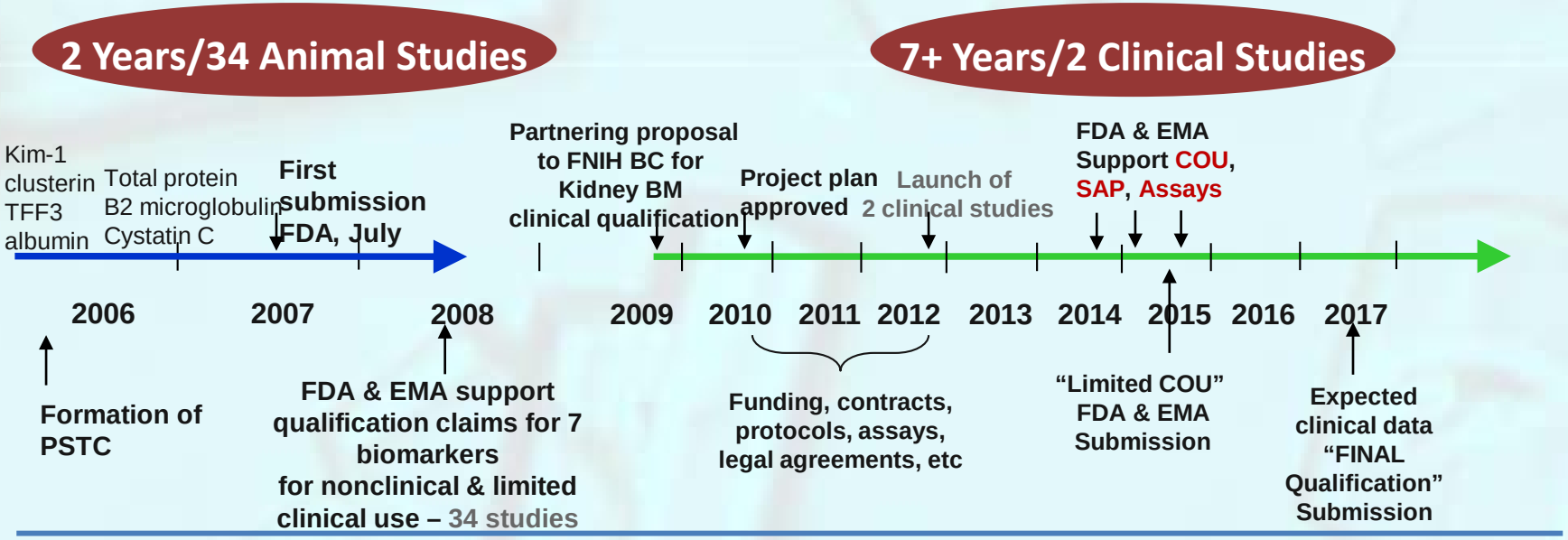


Figure 2. Time course of serum creatinine (S Creat) and urinary kidney damage marker levels. Arrows denote administration of SPC5001 on study days 1, 8, and 15. Conversion factor for S Creat in mg/dL to μmol/L, $\times$88.4. Abbreviations: Ur β2-MG, urinary β_2-microglobulin; Ur GST-a, urinary α-glutathione *S*-transferase; Ur KIM-1, urinary kidney injury molecule 1; Ur NAG, urinary *N*-acetyl-β-D-glucosaminidase.

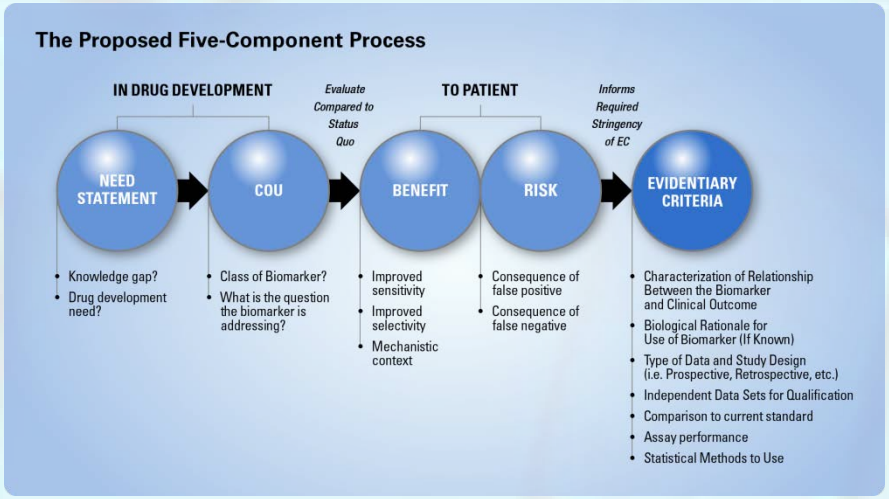
Nonclinical and clinical qualification initiative: timelines and milestones

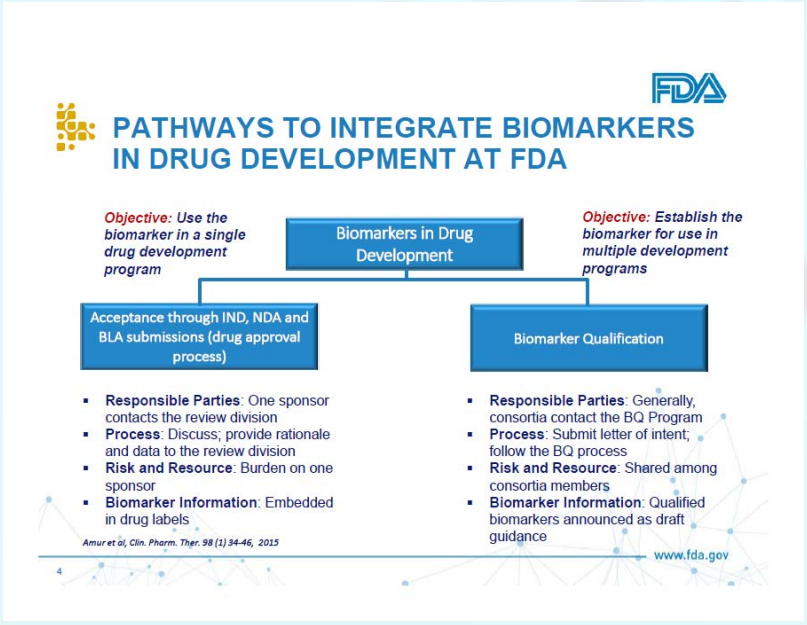


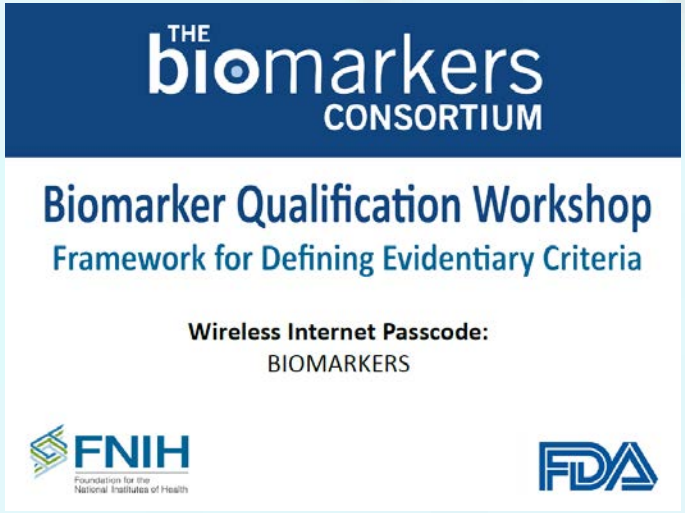
Summary of hypothetical but reasonable examples of drug development scenarios that support the patient health, scientific and business case for qualifying new translational safety biomarkers. [Sistare, Frank D and DeGeorge Joseph J, Biomarkers Med 2011 5(4) 497-514]

Phase of Development	Example	Summary Description	Estimated Benefit from Deploying New Safety Biomarker
Preclinical GLP Animal Toxicology Studies and / or Clinical Trials	#5 Rat-only Kidney Pathology First Seen in Chronic Study	New translational kidney biomarkers demonstrate monitorability of kidney toxicity. Shorter rat studies and chronic monkey studies are negative. Clinical studies show no changes in kidney biomarkers.	Ambiguities about human safety concerns are eliminated. $31M+ in clinical development preserved. Delays in development avoided.
	#6 Dog-only Kidney Pathology Seen in First GLP Study	New translational kidney biomarkers demonstrate monitorability of kidney toxicity seen only in Dog w "medium" margin. Clinical studies conducted show no changes in kidney biomarkers.	Ambiguities about human safety concerns are eliminated. $10M+ in preclinical development preserved. Delays in development avoided.

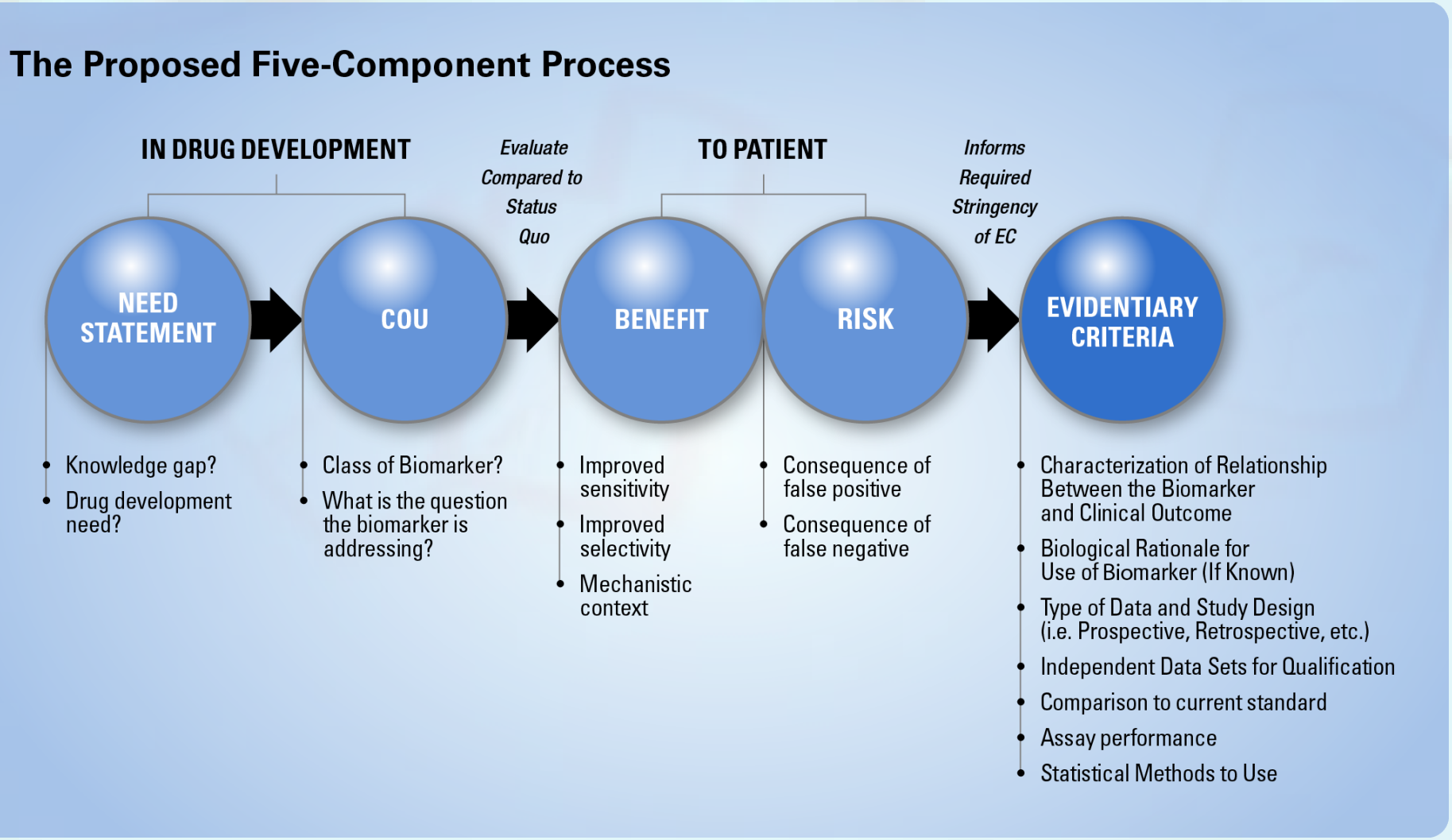
Biomarker qualification: Clarity, predictability, harmonization



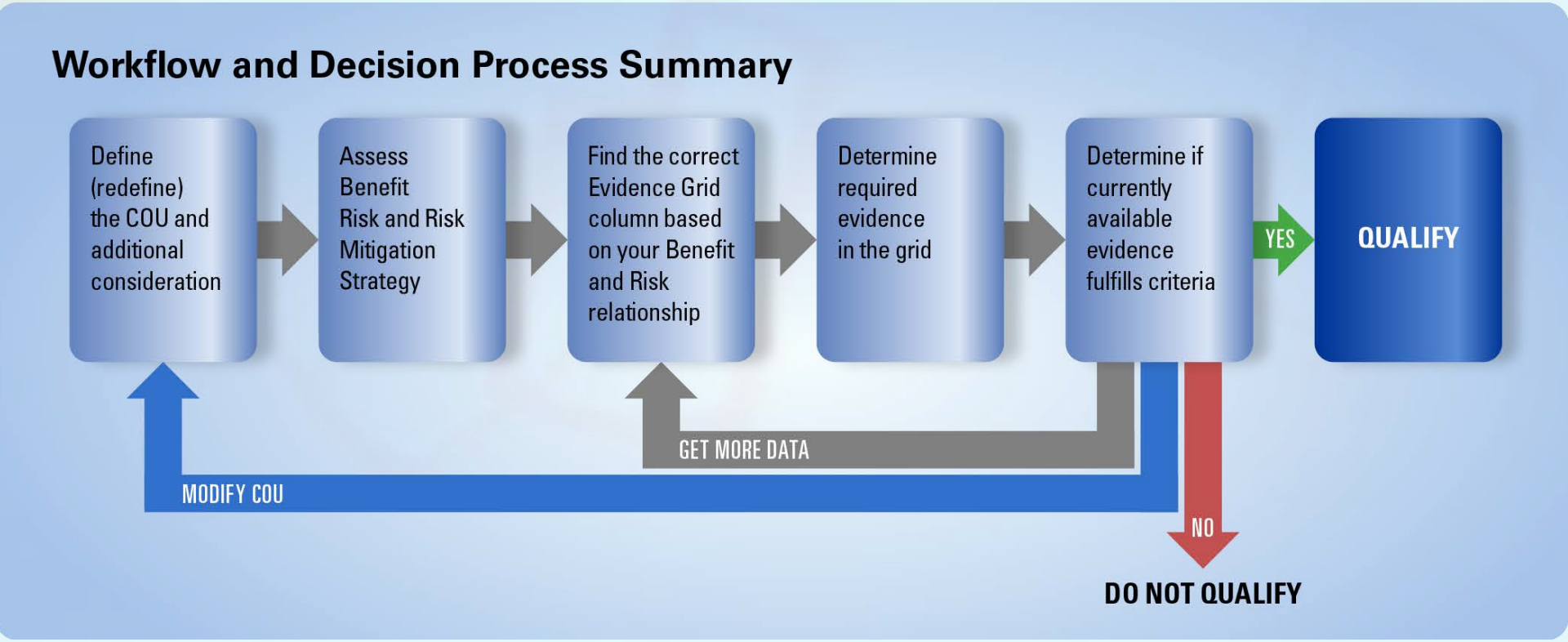




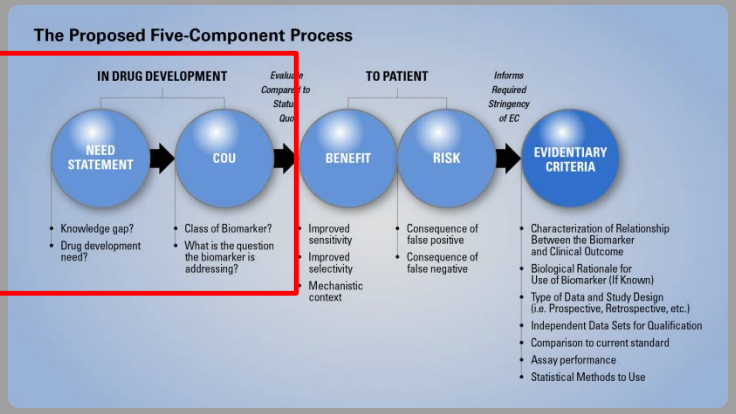
Constructing a biomarker development road map

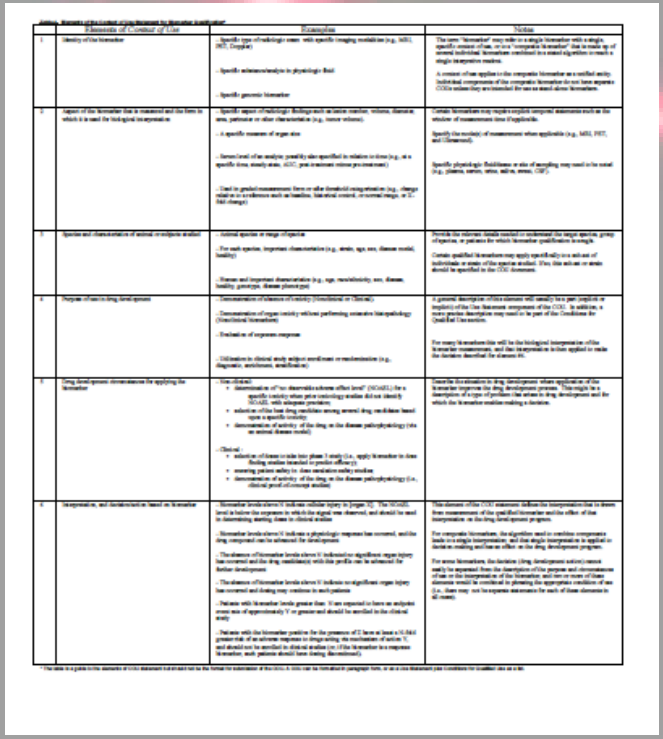


A collaborative approach for biomarker development



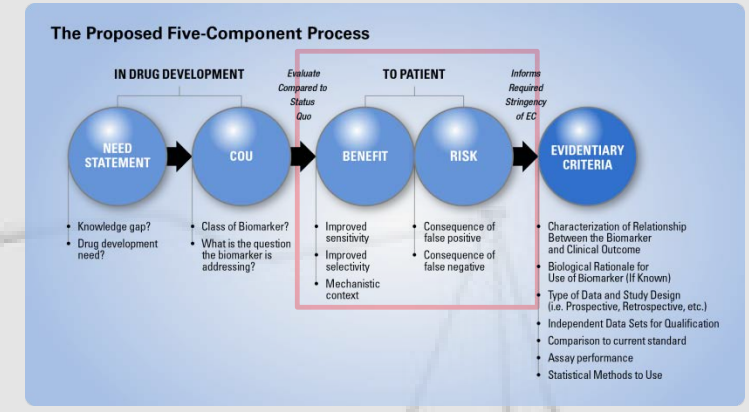
Statement of need and context of use





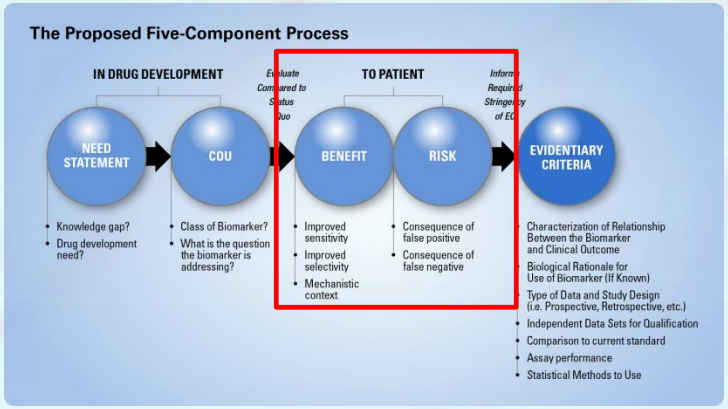
- Statement of need
  - For a biomarker developer and FDA to commit resources for a given project, the need must:
  - direct relevance to drug development
  - potential broad impact
- COU statement – concise description of how a biomarker is intended to be used in drug development
- COU simplified to only 2 elements:
  - What class of biomarker is proposed and what information content would it provide?
  - What question is the biomarker intended to address? (“What is the biomarker’s specific fit-for-purpose use?”)

Benefit and risk



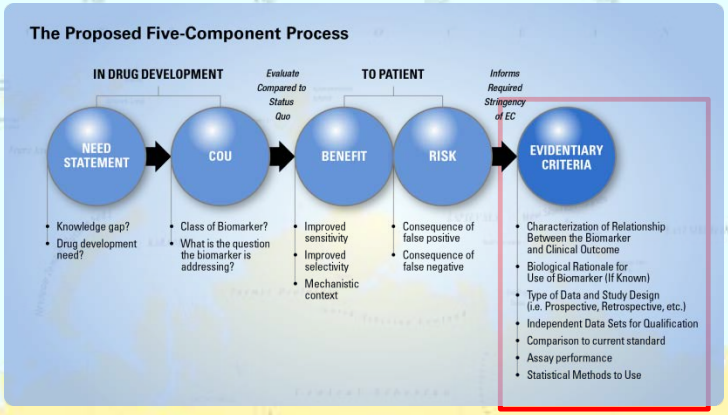
- The benefit and risk profile, given that the COU is related to the biomarker's value to drug development or clinical trials, is assessed from the perspective of patients
- What is the potential consequence or harm if the biomarker performance is not aligned with expectations based on the COU?
  - Ability of a clinical trial to yield interpretable results,
  - Impact on patients enrolled in a clinical trial
  - Impact on patients from a public health point of view should a product be approved or denied approval based, in full or in part, on biomarker information

Examples of benefit and risk analyses



- Favorable benefit and risk profile – lower level of evidence
  - stratification of patients to ensure equal distribution of biomarker positive and biomarker negative individuals in the different arms of a clinical trial
  - If biomarker does not perform – loss of resources but not patient safety
  - But in the setting of a targeted therapies hypothesis testing, more critical
- Less favorable benefit and risk profile – moderate level of evidence
  - Safety biomarker used in addition to the traditional safety biomarkers
  - Degree of risk depends on the impact on decision-making in drug development and the risk to patients enrolled in the trials
- Challenging benefit and risk profile – higher level of evidence
  - Surrogate endpoint
  - If the biomarker is not truly a surrogate endpoint for predicting clinical benefit, results invalid and inappropriate approval decisions made
  - Leads to potentially ineffective drugs marketed or patients denied access to effective therapy

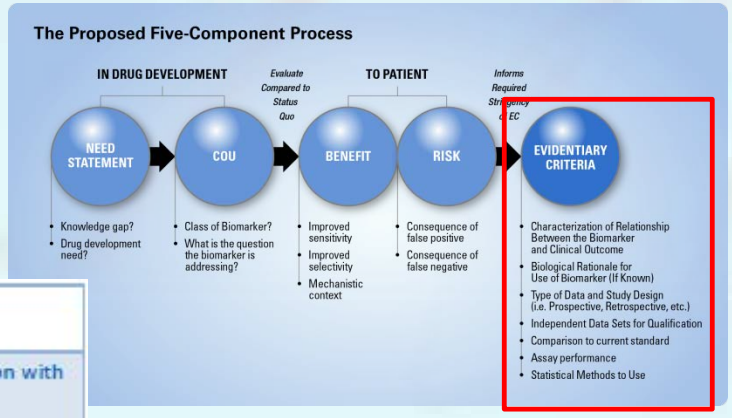
Evidence map



- The evidence maps in this framework are inspired by, but not identical to, the one used by Altar et al. (2008)
- The COU choices made determine the overall relative level of benefit and risk
- Benefit and risk determined as a result of the COU in turn determines the levels of evidence needed to evaluate the biomarker for qualification
- The evidence acceptable for satisfying evidentiary criteria in some cases may be partially or entirely composed of retrospective, literature, or other “real world” types of evidence
- The levels of evidence required to qualify the marker can be described according to a series of variables

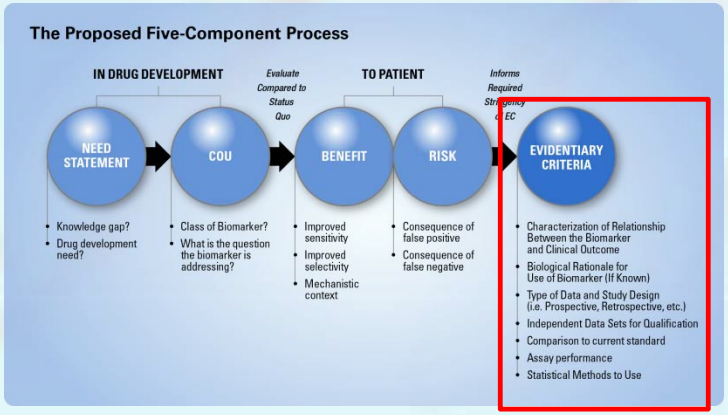
Altar et al. CPT, 83:368-371, 2008

Evidence map



Criterion	High	⟺	Minimal
(1) Assay[1]	Regulatory clearance or approval for marketing as a diagnostic	⟺	"Fit-for-purpose" validation with acceptable performance characteristics
(2a) Scientific Understanding[2]	Causal biological links established between the disease, the intervention and the biomarker	⟺	Gaps in causal links and/or analyte identity
(2b) Scientific Understanding: Data source for comparison of disease to marker	Well designed with focused analysis on one or a small number of biomarkers	⟺	Biomarker discovery analysis from an exploratory trial or dataset
(3) Biological Performance Expectations[3]	Low potential for false result	⟺	Improved performance over current state: [e.g., current standard if available]
(4) Types of data and samples proposed to establish qualification	Prospective double-blind control study or confirmed results in multiple independent data sets	⟺	Retrospective analysis of published results
(4a) Quality of clinical data source: Prospective study	Focused, randomized appropriately powered trial	⟺	Narrow subgroup of intended population, small, or exploratory trial with multiple measures and lack of correction for multiple comparisons
(4b) Quality of clinical data source: Retrospective study	Large population, well controlled combined/meta analysis or multiple studies independently confirming results	⟺	Small, or exploratory trial with multiple measure that is not appropriately powered for significance
(5a)[4] Statistical evidence of the relationship of the biomarker to clinical outcomes	Conclusive across multiple studies	⟺	Some evidence in the literature
(5b) Statistical evidence on the usefulness of the biomarker threshold for significance	Significantly better than current standard (could be in combination with the current standard)	⟺	Similar or slightly better than current standard

Analytical validation



- Accuracy
- Precision
- Analytical sensitivity
- Analytical specificity
- Reportable range
- Reference interval
- Reproducibility
- Stability

Very High Standard:	**Minimum Requirements:**
Regulatory Marketing Approval as Diagnostic	"Fit-for-Purpose" Validation
Parameters Evaluated During Validation	
Accuracy Precision Analytical sensitivity Analytical specificity Reportable range Reference interval Reproducibility Stability Other as required	Accuracy Precision Analytical sensitivity Analytical specificity Reportable range Reference interval Other as required

Conclusion

- Alignment from multiple, diverse stakeholders
- Consistent, comprehensive, semi-quantitative parameters for biomarker qualification
- Greater degree of clarity, predictability, and harmonization
- Broadly applicable across multiple categories of biomarkers and COUs
- Since each category of biomarker and COU has unique factors to consider as part of the development process, multiple modules are proposed to address these more specific issues

Thanks to .com, .edu, .gov, and.org!

- **Evidentiarly Criteria Working Group**
  - Linda Brady, NIMH/NIH
  - Martha Brumfield, C-PATH
  - Bill Chin, PhRMA
  - Steve Hoffmann, FNIH
  - Gary Kelloff, NCI/NIH
  - Gabriela Lavezzari, Duke
  - Chris Leptak, FDA
  - Joe Menetski, FNIH
  - Rajesh Ranganathan, PhRMA
  - John-Michael Sauer, C-PATH
  - Frank Sistare, Merck
  - John Wagner, Takeda
  - David Wholley, FNIH

- **Drug Induced Kidney Injury Lead**
  - Frank Sistare, Merck
  - Steve Hoffmann, FNIH

- **Analytical Validation Team**
  - Amanda Baker, C-PATH
  - Steven Piccoli, BMS
  - John-Michael Sauer, C-PATH
  - Diane Stephenson, C-PATH

- **Statistical Team**
  - Aloka Chakravarty, FDA
  - Suzanne Hendrix, Pentara
  - Lisa McShane, NCI/NIH
  - Robin Mogg, Merck
  - Klaus Romero, C-PATH
  - Sue Jane Wan, FDA

- **Drug Induced Liver Injury Lead**
  - Jiri Aubrecht, Pfizer

- **Drug Induced Vascular Injury Lead**
  - Brad Enerson, Pfizer
  - Michael Lawton, Pfizer
  - Tanja Zabka, Genentech