

Evidentiary Considerations for Integration of Biomarkers in Drug Development : Statistical Considerations

August 21, 2015

Aloka Chakravarty, PhD

Office of Biostatistics, OTS, CDER
U.S. Food and Drug Administration

Aloka.Chakravarty@fda.hhs.gov

All views are those of the speaker and should not be construed as the official FDA Policy

Outline

- Biomarker – definition and its challenges
- Types of biomarkers and their role in drug development
- Evidentiary Standards – Statistical Design Issues
 - Identification
 - Biomarker levels and Threshold
 - Reference standards
- Evidentiary Standards – Statistical Analysis Issues
 - Cross-validation
 - Interim Analysis
 - Analysis plan considerations
 - Prospective-retrospective analysis issues
- Conclusions

Biomarker – Definitions and its challenges

- A **biomarker** is a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes (abnormal biologic processes) or responses to a therapeutic intervention. It is not a clinical assessment of the patient, those evaluating or closely relating to how a patient feels or functions, or survival (Biomarker Definition WG, 2001).
- The committee observed a great deal of *inconsistent and imprecise definition* and use of terms relevant to biomarkers and biomarker evaluation. Consistent, precise definition and use of terms is critical Institute of Medicine (US) Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease; Micheel CM, Ball JR, editors. Washington (DC): [National Academies Press \(US\)](#); 2010
- A FDA-NIH Joint Council Working Group has started work on providing a consistent definition of biomarkers - FDA, NIH and NLM

Types of biomarkers and their role in drug development

Biomarker Use	Drug Development Objective
Disease risk stratification	Assess the likelihood that the disease will develop (or recur)
Prevention	Identify and track risk factors
Screening	Detect and treat early-stage disease in the asymptomatic population
Diagnosis	Definitively establish the presence of disease
Classification	Classify patients by disease subset
Prognosis	Predict the probable outcome of disease to determine the aggressiveness of treatment
Prediction/treatment stratification	Predict response to particular therapies and choose drug mostly likely to yield favorable response (outcome) in patient
Therapy-related risk management	Identify patients with a high probability of adverse effects of a treatment
Therapy monitoring	Determine whether a therapy is having the intended effect on a disease and whether adverse effects arise
Surveillance	Early detection (and treatment) of advancing disease or complications

Adapted from IOM Biomarkers & Surrogates Report, 2010

Biomarker Uses in Drug Development



Biomarker Use	Drug Development Objective
Target validation	Demonstrate that a potential drug target plays a key role in the disease process
Early compound screening	Identify compounds with the most promise for efficacy and safety
Pharmacodynamic assays	Determine drug activity; select dose and schedule
Patient selection	In clinical trials, patient selection (inclusion/exclusion)
Surrogate endpoint	Use of an alternative outcome measure which can be measured sooner, less invasively, or with less inconvenience or cost, in place of the long-term primary endpoint to determine more quickly whether the treatment is efficacious and safe in drug regulatory approval

Adapted from IOM Biomarkers & Surrogates Report, 2010

- Enrichment¹: “prospective use of any patient characteristic to select a study population in which detection of a drug effect (if one is in fact present) is more likely than it would be in an unselected population”
 - Strategies to decrease heterogeneity – reduce inter-patient and intra-patient heterogeneity
 - Prognostic enrichment strategies – choosing patients with a greater likelihood of having a disease-related endpoint event
 - Predictive enrichment strategies – choosing patients more likely to respond to the drug treatment (i.e., treatment selection biomarker)

<http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm332181.pdf>

Evidentiary Standards

Pre-considerations

- Relation of the biomarker to its Context of Use (CoU)
 - acceptable, analytically validated measurement method
 - the operating characteristics of the assay, including normal ranges and distribution of values and variance
 - for composite, relationship of the biomarker's components to each other
 - Patient characteristics or covariates that have an effect on biomarker expression
- Learn and confirm paradigm
 - Exploratory analysis in learning stage: Expression, threshold and relevant covariates
 - Confirmatory: sample size, H_0 testing, analysis plans

Evidentiary Standards – Design Issues

- Identification
 - methods for identifying multiple predictors: regression with multiple covariates, ridge regression, LASSO
 - Tree-based methods: non-linear relationships
- Biomarker Levels
 - Baseline measure?
 - Change relative to a baseline? Difference or relative change?
 - If intra-subject variability $>$ inter-subject variability \rightarrow linear change from baseline better; relative change from baseline is preferred when the opposite holds.

Evidentiary Standards – Design Issues (2)

- Threshold

- Definition can be complicated; may involve single or multiple time points - care taken to establish relevance for each component (e.g., disease severity, intended population)
- Replicability important - mitigated if a large number of patients highly representative of the population studied
- Model selected to investigate the threshold should be as parsimonious - avoid over-fitting
- Multiplicity may also be an issue

Evidentiary Standards – Design Issues (3)

- Reference standards
 - If no established gold standard biomarker, use all available information
 - If a flawed gold standard (call it the “pseudo-gold standard”) is used as reference,
 - the new biomarker may lack sensitivity
 - its estimation may be biased - extent of this bias depends on the correlation between the new and pseudo-gold standard biomarker
 - Use of adjudication committees may mitigate this bias

Evidentiary Standards – Analysis Issues

- Cross validation (CV)
 - Training and validation set (single-fold validation) or carried forward in a k-fold validation?
 - Is validation more credible coming from cross validation applied to a single dataset or from a separate trial?
 - If credible validation comes from a separate trial, there should be no prior knowledge of the outcomes before the biomarker evaluation
 - Trade-off between bias and variance
 - adequate sample size is needed in each training set to provide an unbiased estimate of the true prediction error
 - If all model building steps are included in CV, with no external variable selection or outcome evaluation, CV can be a powerful tool

Evidentiary Standards – Analysis Issues (2)

- Interim Analysis
 - Can be planned in both learning and confirming phases
 - Earlier interim looks
 - to assess the initial performance
 - modify biomarker thresholds or for sample size re-estimation
 - Later interim looks
 - assess the performance improvements based on the modifications, including thresholds
 - stop the trial based on futility
 - For each interim look
 - Clearly identify its objective, any associated sample size re-estimation and the effect of the interim analysis on Type I error.
 - If a limited COU is initially planned, it can guide future development plans based on the learning phase information.

Evidentiary Standards – Analysis Issues (3)

- Analysis Plan
 - Each hypothesis of interest has to be pre-specified, along with its relevant analysis plan.
 - Multiplicity adjustments
 - Procedures to handle missing data
 - Plans for any secondary comparison
 - If any confirmatory subgroup analysis is planned, it needs to be pre-specified with a cross-validation or hierarchical testing strategy to avoid inflation of experiment-wise Type I error

Evidentiary Standards – Analysis Issues (4)

- Key elements for retrospective analysis to be adequate
 - Studies adequate, well controlled, have large enough sample size to ensure adequate power
 - Prognostic factors balanced across treatments in subgroups
 - Biomarker reflects ITT population – no convenience sample
 - Assay well-characterized, acceptable analytical performance - same assay used in all the studies.
 - Integrity in question if analysis plan occurs after the efficacy data has been unblinded and the biomarker status known
 - Analysis plan controls multiplicity and the study-wise Type I error
 - Retrospective evaluation not used to salvage a negative study

Evidentiary Standards – Analysis Issues (5)

- Prospective-retrospective design
 - The biomarker hypothesis is *prospectively specified* prior to diagnostic assay testing.
 - Samples are collected prior to treatment initiation and may be stored for later use. The biomarker classification is then conducted using a validated assay to characterize the biomarker for the proposed COU.
 - The clinical outcome data may have already been (partially) collected, unblinded, and analyzed. However, if the prior analyses did not include biomarker data, the biomarker analysis might be considered as “prospectively” performed with a “retrospective classifier analysis”.

Concluding Remarks

- Meticulous planning prior to undertaking the project critical
- Account for mid-course modifications
- Collaboration across stakeholders important
- Multiregional factors can be challenging – plan for it
- Early engagement of regulators can improve chances of overall success

Acknowledgements:

Xavier Benain

Lisa LaVange

Lisa McShane

Robin Mogg

Matthew Schipper

Sue Jane Wang

Questions?