Key Points for IVD Clinical Studies

Abraham Tzou, MD

Proteomics in the Clinic

June 13, 2014

- Design matters
- Support claims
- Study fundamentals
  - Beware of bias!
- Study interpretation

Study Design Matters

- A flawed test should fail in analytical and/or clinical validation because it fundamentally doesn't work. This failure is a *good* thing.
- A safe and effective test may fail in analytical and/or clinical validation because of improper study design. This failure is a *bad* thing.

➢ Check the basics in IVD clinical studies

Research and discovery deal with means and populations, but medicine deals with individual patients. Keep in mind how the study results will apply to the patient *when used clinically*.

Issue #1: Scope of Claim(s)

The assay is an in vitro diagnostic test designed to *diagnose*, *monitor* breast cancer and *predict* response to treatment and provide *prognosis*.

Intended Use Statement (how/by whom device is used)

- What is the device measuring, identifying or detecting? (analyte, organism, ..)
- Specimen types, sources (whole blood, serum,..)
- Conditions for use (hospital lab, home use,..)
- What type of data output? (quantitative, qualitative, semi-quantitative)

Indication for Use Statement (for what/on whom device is used)

- Target condition
  a particular disease, a disease stage, health status, or any other identifiable condition of event within a patient
- Target population (intended use population)
  those subjects for whom the test is intended to be used
- Medical Testing Contexts
  as, for screening, diagnosis, monitoring, prognosis and so on.

Different Contexts → Different Claims

- **Diagnosis** (target condition is present or not during the time of testing)
- **Screening** (maybe in a general population [asymptomatic subjects at average risk])
- **Risk** (assessment of predisposition to disease in future)
- **Prognosis** (separating already diagnosed patients into poor or good outcomes)
- **Monitoring** (evaluating change in a patient's condition)
- **Companion Diagnostics/Co-development paradigm** (whether a patient is a candidate for therapy)

This is not a comprehensive list

➢ Clinical study design stems from Intended Use/Indications for Use claims (e.g., cross-sectional vs. longitudinal data)

Different Claims → Different Designs

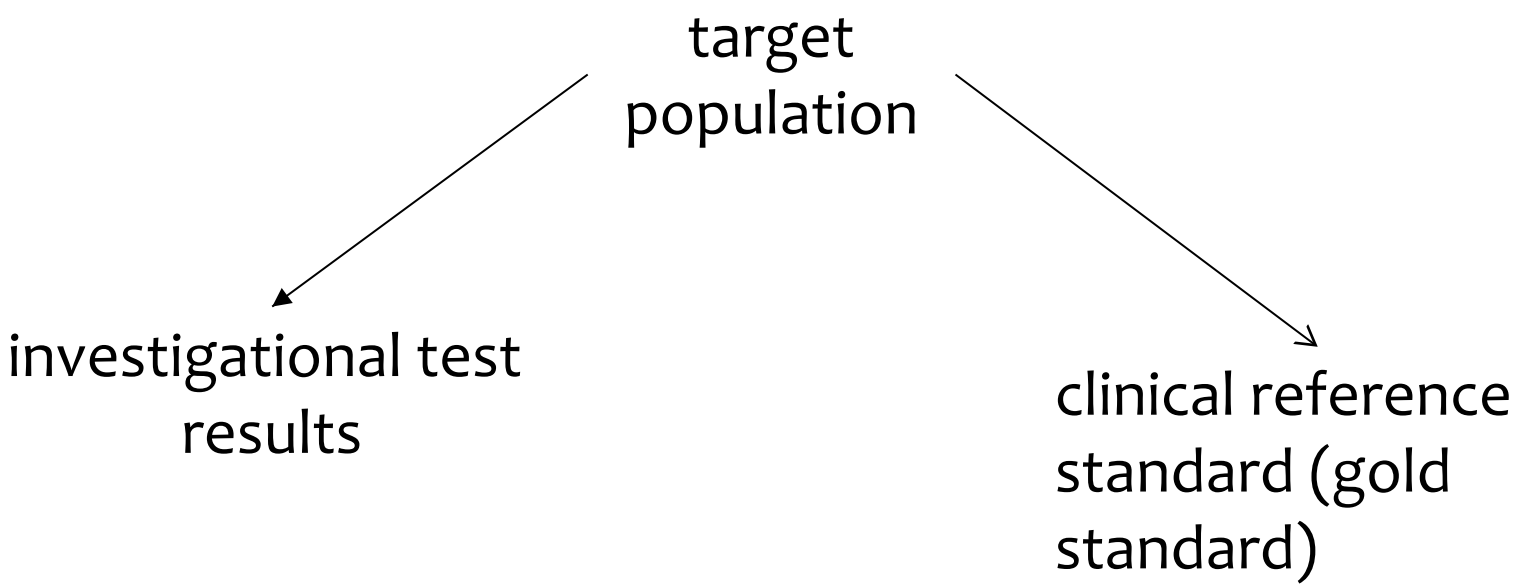
- Screening for cancer:
  - Low prevalence in general population -> large cross-sectional study
- Monitoring for cancer
  - Already diagnosed patients followed over time, longitudinal study with pre-specified definition of *clinically relevant* change

➢ Evidence required to support multiple claims may come from separate studies.

Issue #2 – Inappropriate Study Design

- 50 prostate cancer positive
- 50 *healthy* blood donors
- 24 *women* with breast cancer
- 5 benign prostatic hyperplasia samples

Study Design Fundamentals



- No technology or sample size fix for flawed study

Beware of Potential Biases

We considered an ideal scenario when N randomly selected subjects are from the intended use population and each subject has result of the test and verification of disease (D+, D-).

Potential Biases
1) Selection bias (when the study population does not represent the intended use population) – spectrum bias
2) Verification bias

Banked (retrospective) samples (potential selection biases)

- How representative are banked samples (inclusion/exclusion criteria)
- Only leftovers from big tumors (sample volumes)?
- Storage does not impact analyte of interest

➢ Provide unbiased estimates of performance

How banked samples can lead to Spectrum Bias

Example

Diseased subjects in the Intended Use population =
50% of Stage II and 50% of Stage I
Test ABC has sensitivity for Stage II = 90%; Stage I = 50%

Sensitivity of test ABC in the IU population =
0.5 * 90%+ 0.5 * 50% = **70%**

Retrospective samples in the clinical study
80% of Stage II and 20% of Stage I:
Sensitivity in the clinical study =0.8 * 90% + 0.2 * 50* = **82%**

Sensitivity is biased (overestimated)

Improper Sample characterization can lead to Verification Bias

Example

Clinical study with 100 subjects: each subject has verification of disease and test result

		Gold Standard		Total
		D+	D-	
Test	Pos	20	5	25
	Neg	30	45	75
Total		50	50	100

Se = 40% (20/50)

Sp = 90% (45/50)

Example (cont.)

Subjects were referred to the Gold Standard (GS) based on the "Current clinical practice".

In the study, all 25 subjects with pos. test results -> GS;
only 1/3 of 75 subjects with neg. test results -> GS.

Analysis of the data with verified disease status

		Gold Standard		Total
		D+	D-	
Test	Pos	20	5	25
	Neg	10	15	25
Total		30	20	50

Se = 67% (20/30) Sensitivity is biased (overestimated)

Sp = 75% (15/20) Specificity is biased (underestimated)

Study Design Fundamentals

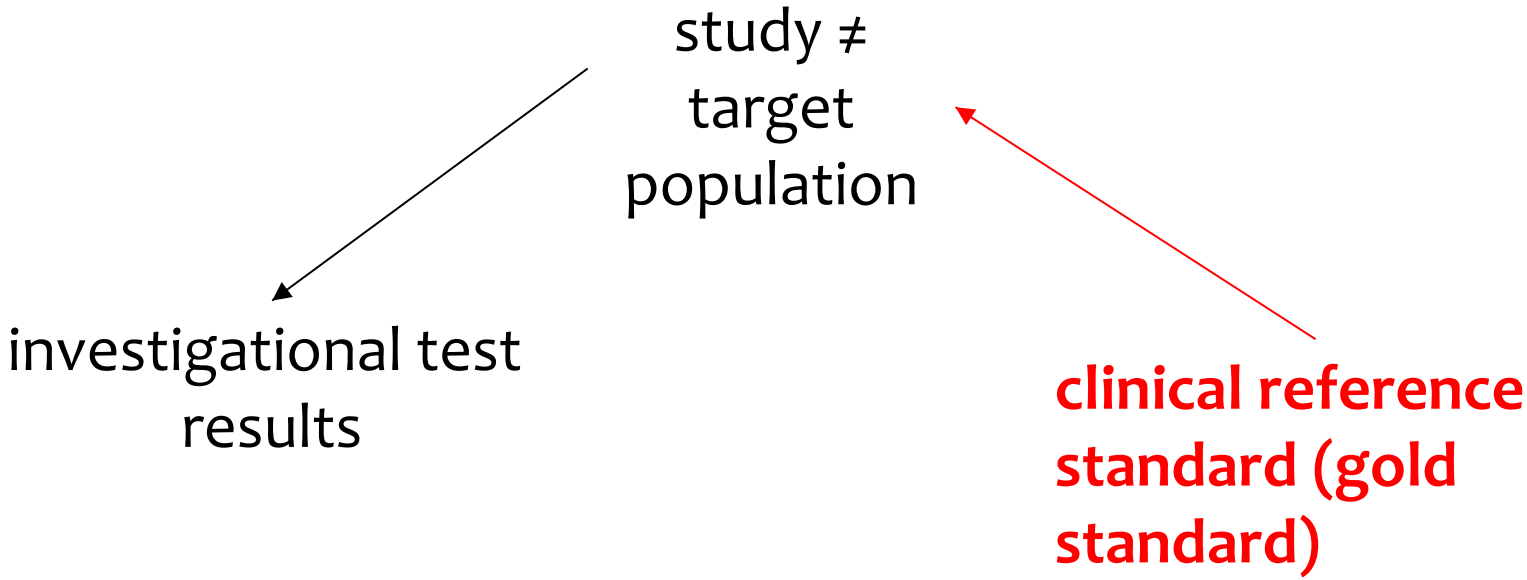
study ≠ target population

investigational test results

clinical reference standard (gold standard)

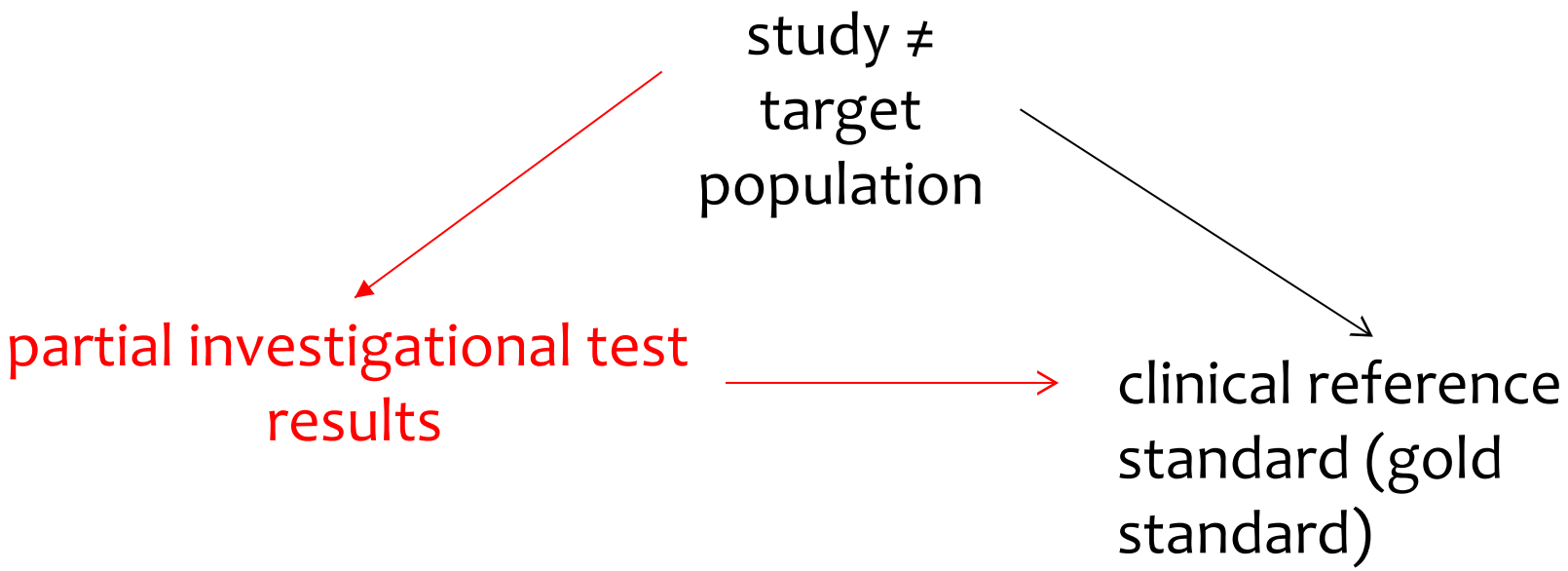
e.g., specialist vs. primary care, high risk vs. general risk, already diagnosed vs. early diagnosis

Study Design Fundamentals



e.g., only include patients undergoing procedure, exclude patients without follow-up

Study Design Fundamentals



e.g., samples unavailable, invalid results, investigational results confound gold standard evaluation, test and/or analysis plan not finalized

Issue #3 – Study Interpretation

The black solid bars within the boxplot represent the median abundance, and the dotted line represents mean abundance for the given group. Error bars represent s.d. *The P values for analytes between groups were P <0.05.*

In patients for whom *biopsy is recommended*, positive results are associated with *an increased likelihood of a positive biopsy*.

Study Interpretation Issues

- Clinically meaningful performance
  - Statistical significance may not be medically important (e.g., no impact on clinical decision)
- Comparison to current practice
  - Benefit risk in relation to other options
- Performance consistency across patients
  - Added value over clinical covariates

HPV approved for cervical cancer screening

- Evaluated target population for the claim
- Addressed study design fundamentals
  - Appropriate sample size for prevalence
  - Unbiased sample selection and characterization
- Clinically meaningful performance
- Benefit-risk compared to clinical practice

- Design matters
- Support claims
- Study fundamentals
  - Beware of bias!
- Study interpretation