



U.S. Food and Drug Administration

Notice: Archived Document

The content in this document is provided on the FDA's website for reference purposes only. It was current when produced, but is no longer maintained and may be outdated.

Diagnostic Devices for Tailoring Therapies

Gene Pennello, Ph.D.
Team Leader, Diagnostics Devices Branch
Division of Biostatistics
Office of Surveillance and Biometrics
Center for Devices and Radiological Health, FDA

Outline

- Tests for tailoring therapies
- Validation of In Vitro Diagnostics
- Trial Designs and Scenarios
- Multiple Testing
- Bridging, with an example of an analysis
- Review Issues

Tests to Select Therapies

- Safety
 - CYP2D6 genotypes' effect on metabolic rate for drugs
 - HLA allele B*1502 as a marker for carbamazepine-induced Stevens-Johnson syndrome and toxic epidermal necrolysis
 - UGT1A1 genotype for risk of neutropenia in CRC patients taking irinotecan
 - KRAS mutation for likely absence of cetuximab, panitumumab efficacy in CRC patients
- Effectiveness
 - HER2 positive breast cancer patient selection for trastuzumab
 - EGFR to select CRC patients for cetuximab, panitumumab.
- Dosing
 - VKORC1 and CYP2C9 genotype to predict warfarin dose.

Scenarios in Drug-Dx Development

| | Diagnostic | |
|------|------------|--|
| Drug | Old | New |
| Old | | |
| New | | Focus of 2005 Drug/Dx Co-development Concept Paper |

Scenarios in Drug-Dx Development

| | Diagnostic | |
|------|---|---|
| Drug | Old | New |
| Old | warfarin:2C9+VKORC1 | irinotecan:UGT1A1 cetuximab:KRAS [†] panitumumab:KRAS [†] |
| New | I-SPY Phase II trials panitumumab:EGFR | trastuzumab:HER2 [†] cetuximab:EGFR |

[†]Bridge from lab developed test (LDT) to market ready test (MR₅T).

In Vitro Diagnostics (IVDs)

Submission Requirements:

- Analytical Validation: *does my test accurately and reproducibly measure the analyte I think it does? Correctly? Reliably?*
- Clinical Validation: *does my test result correlate with the expected clinical presentation? How reliably?*

Analytical Validation

- Analytical performance
 - Precision (repeatability, reproducibility)
 - Accuracy
 - Sensitivity (limit of detection)
 - Specificity (interference, cross-reactivity)
 - Sample type / matrix
 - Sample preparation / conditions
 - Performance around the cut-off
 - Potential for carryover, cross-hybridization

Clinical Laboratory Standards Institute

- FDA recognizes several CLSI guidelines:
 - **EP5** Precision Performance of Quantitative Measurement Methods
 - **EP6** Linearity of Quantitative Measurement Procedures
 - **EP9** Method Comparison and Bias Estimation Using Patient Samples
 - **EP12** Qualitative Test Performance
 - **EP14** Evaluation of Matrix Effects
 - **EP17** Limit of Detection

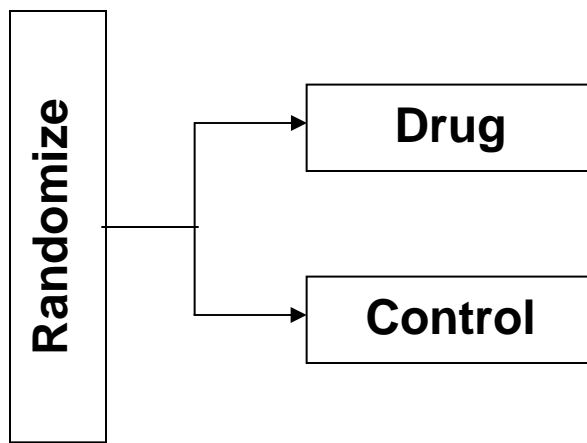
Clinical Validation

- Cytochrome P450 2D6 genotyping assay
 - o literature validation was acceptable
- Novel gene or a multi-gene classifier for predicting efficacy for a specific drug
 - o most likely needs clinical validation

Trial Designs

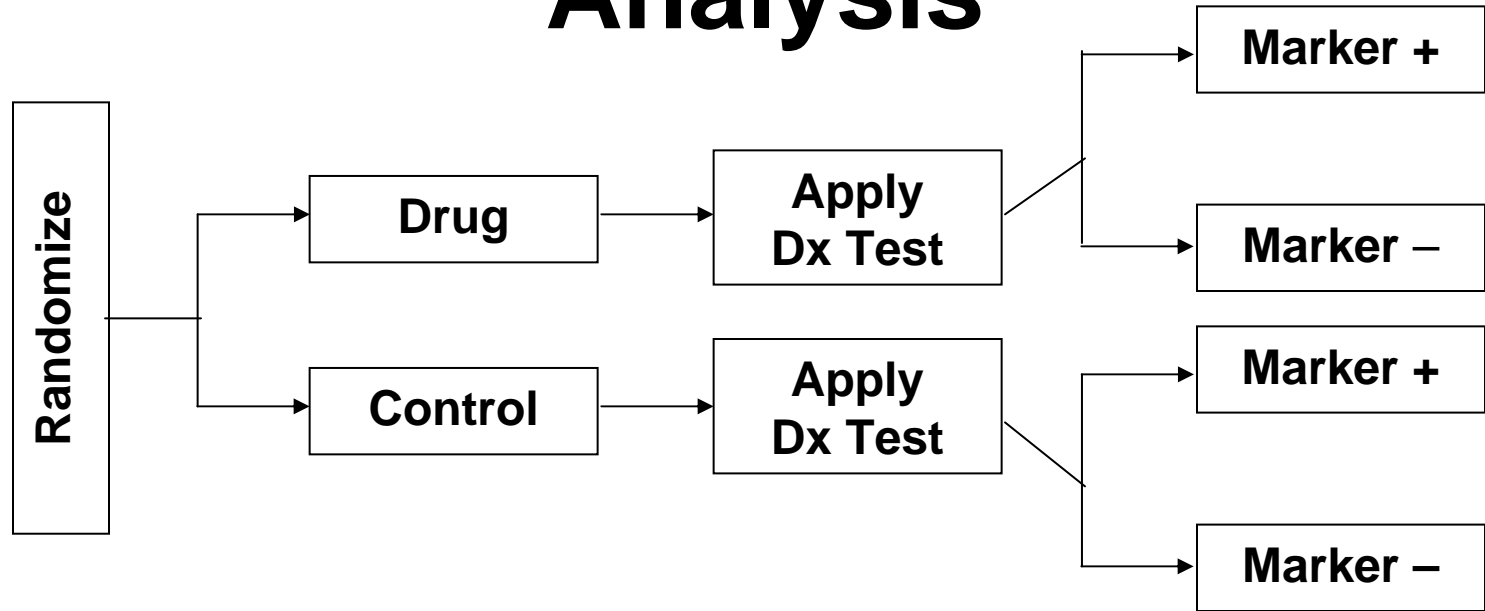
- Reference Design
- Marker by Treatment Design
 - Randomization is stratified by marker
- Targeted Design
 - Enroll a subgroup defined by marker.

Reference Design



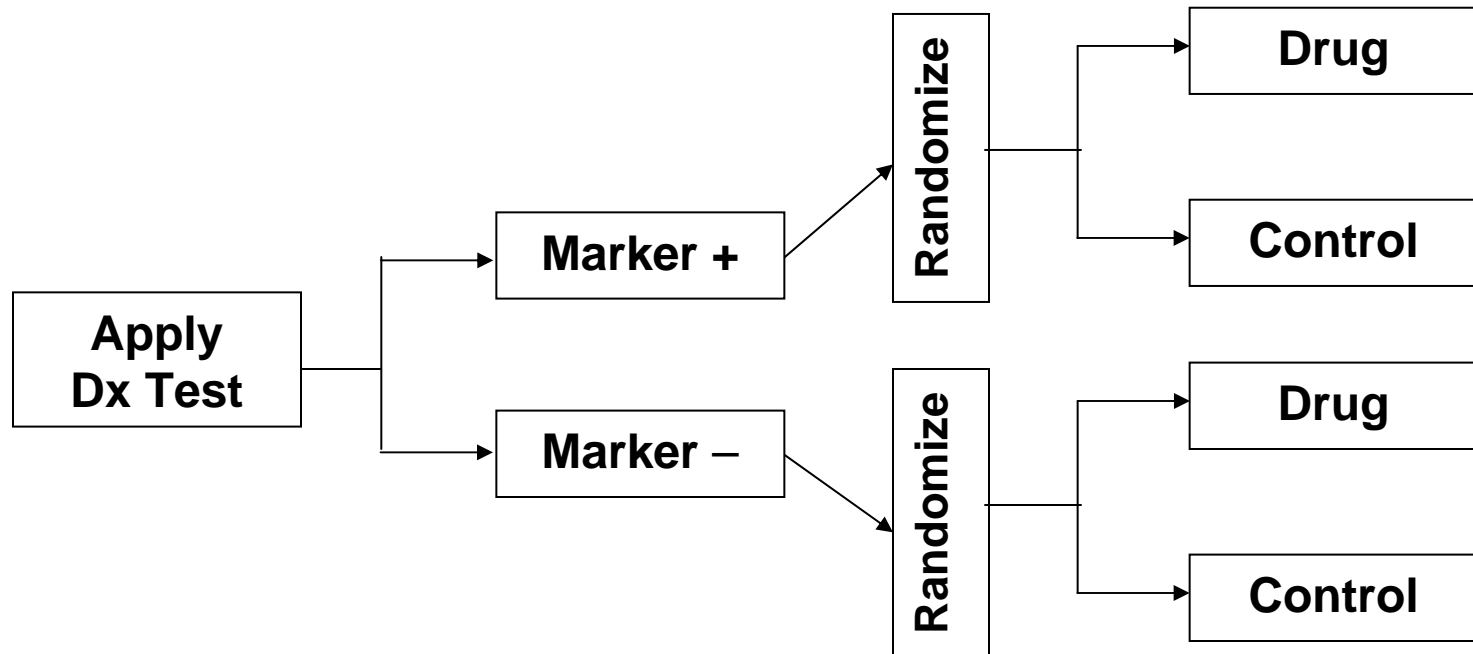
- Evaluate if drug works overall.
 - Mixes marker + and – drug effects.
- Save samples for prospective, retrospective analysis of a marker-defined subset.

Prospective-Retrospective Analysis



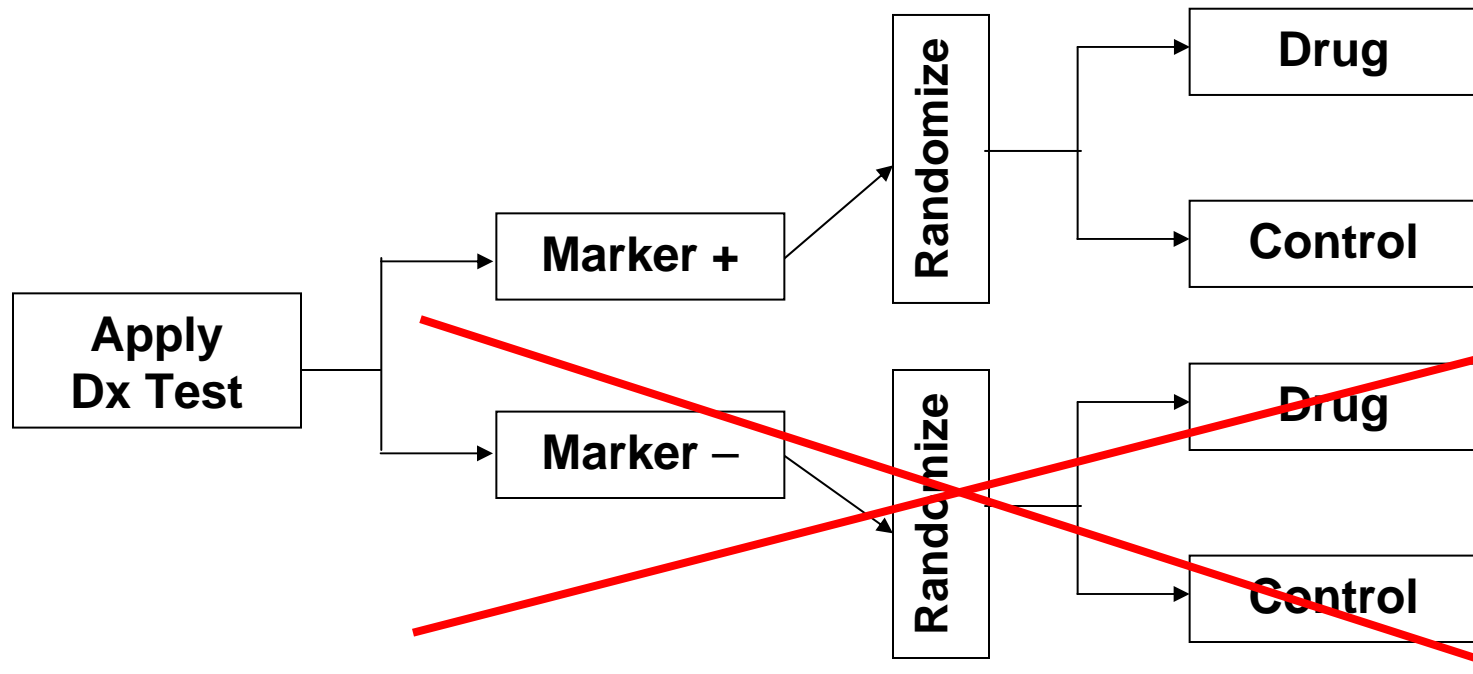
- Dx result may be missing for some patients.
 - Insufficient sample
 - Lack of consent
- If not missing at random, then results biased.

Marker by Treatment Design



- Dx test must be available at time of study.
- Improves covariate balance within subgroups.
- Opportunity to enrich low prevalence subset₃

Targeted Design



- Can be efficient for studying drug efficacy.
- Assumes biology is well-understood
- Test is not studied for effectiveness.
 - Test approval, reimbursement problematic.

Drug vs. Marker Effectiveness

- **Drug efficacy**
 - in overall population, or
 - in a subgroup defined by the test
- **Device effectiveness**
 - marker predicts a differential effect of drug (drug by device interaction)
 - Cannot assess effectiveness in targeted trials.

Multiple Testing for Drug Efficacy

1) Single Step Procedure (Budget alpha)

- Test overall effect at 0.04
- Test marker-defined subgroup at 0.01
- Subgroup effect may drive overall effect.
- Interpretation could be difficult.

Multiple Testing for Drug Efficacy

2) Step-Down Testing

- Test null for subgroup at 0.05.
 - If reject, get subgroup claim, else STOP.
- Test null for complement at 0.05.
 - If reject, get overall claim.
- Easier to interpret.
- Not impractical if subgroup is not rare or you enrich for it.

Bridging

- Market ready test (MRT) often unavailable at time of Phase III trial.
- Instead, a laboratory developed test (LDT) is used in conduct of trial
 - Stratify randomization by marker (MbT design)
 - Enroll marker + patients (targeted design)
- Bridge is made from LDT to MRT, when it becomes available, using:
 - available samples from Phase III, or
 - external samples



Limitations of Bridging

- Primary analysis for drug based on LDT
 - What if re-analysis using MRT results provides different conclusions?
- Concordance between MRT and LDT
 - If MRT and LDT are not 100% concordant, then can't tell if drug efficacy is worse or better in MRT+ subgroup than LDT+ subgroup.
 - What level of concordance is high enough?
- Targeted trials
 - Concordance only on LDT +'s, not LDT – 's.
 - Drug efficacy in MRT + conditional on LDT+

Bridging: MRT_1 to MRT_2

| <i><u>Study</u></i> | <i><u>MRT_1</u></i> | <i><u>MRT_2</u></i> | <i><u>Drug</u></i> |
|---------------------------|----------------------------------|----------------------------------|--------------------|
| <i>Phase III</i> | <i>X</i> | | <i>X</i> |
| <i><u>Concordance</u></i> | <i>X</i> | <i>X</i> | |

Goal: Estimate drug efficacy in subgroup defined by MRT_2 .

Hypothetical Data

- **Phase III Data**

- Reference or Marker by Treatment Design
- Objective Response Rate, Drug Arm Only

| | Response | | | Response Rate |
|------------------|----------|-----|-------|------------------|
| MRT ₁ | No | Yes | Total | |
| – | 109 | 67 | 176 | 38.1% |
| + | 284 | 333 | 617 | 53.9% |
| Total | 393 | 400 | 793 | |



Hypothetical Data

- **Concordance Data**

- External samples
- Population assumed same as in Phase III.
- Simple random sample (no enrichment).

| | MRT ₂ | | | Agreement |
|------------------|------------------|-----|-------|-----------|
| MRT ₁ | – | + | Total | |
| – | 70 | 19 | 89 | 78.7% |
| + | 66 | 245 | 311 | 78.7% |
| Total | 136 | 264 | 400 | |

Phase III Data

- Brackets denote missing data.

| Non-Responders | | | |
|------------------|------------------|-----------------|-------------------|
| | MRT ₂ | | |
| MRT ₁ | – | + | Total |
| – | $[x_{000}]$ | $[x_{010}]$ | $x_{0\cdot0}$ |
| + | $[x_{100}]$ | $[x_{110}]$ | $x_{1\cdot0}$ |
| Total | $[x_{\cdot00}]$ | $[x_{\cdot10}]$ | $x_{\cdot\cdot0}$ |

| Responders | | | |
|------------------|------------------|-----------------|-------------------|
| | MRT ₂ | | |
| MRT ₁ | – | + | Total |
| – | $[x_{001}]$ | $[x_{011}]$ | $x_{0\cdot1}$ |
| + | $[x_{101}]$ | $[x_{111}]$ | $x_{1\cdot1}$ |
| Total | $[x_{\cdot01}]$ | $[x_{\cdot11}]$ | $x_{\cdot\cdot1}$ |

Concordance Data

- Brackets denote missing data.

| MRT ₂ – | | | |
|--------------------|------------------|------------------|--------------------|
| | Response | | |
| MRT ₁ | No | Yes | Total |
| – | $[y_{000}]$ | $[y_{001}]$ | $y_{00\cdot}$ |
| + | $[y_{100}]$ | $[y_{101}]$ | $y_{10\cdot}$ |
| Total | $[y_{\cdot 00}]$ | $[y_{\cdot 01}]$ | $y_{\cdot 0\cdot}$ |

| MRT ₂ + | | | |
|--------------------|------------------|------------------|--------------------|
| | Response | | |
| MRT ₁ | No | Yes | Total |
| – | $[y_{010}]$ | $[y_{011}]$ | $y_{01\cdot}$ |
| + | $[y_{110}]$ | $[y_{111}]$ | $y_{11\cdot}$ |
| Total | $[y_{\cdot 10}]$ | $[y_{\cdot 11}]$ | $y_{\cdot 1\cdot}$ |

Bayesian Analysis

- Underlying probabilities

| Non-Responders | | | |
|------------------|------------------|---------------|-------------------|
| | MRT ₂ | | |
| MRT ₁ | – | + | Total |
| – | p_{000} | p_{010} | $p_{0\cdot0}$ |
| + | p_{100} | p_{110} | $p_{1\cdot0}$ |
| Total | $p_{\cdot00}$ | $p_{\cdot10}$ | $p_{\cdot\cdot0}$ |

| Responders | | | |
|------------------|------------------|---------------|-------------------|
| | MRT ₂ | | |
| MRT ₁ | – | + | Total |
| – | p_{001} | p_{011} | $p_{0\cdot1}$ |
| + | p_{101} | p_{111} | $p_{1\cdot1}$ |
| Total | $p_{\cdot01}$ | $p_{\cdot11}$ | $p_{\cdot\cdot1}$ |

Bayesian Analysis

- **Prior Distribution (non-informative)**

$$\underline{p} = (p_{000}, p_{010}, p_{100}, p_{110}, p_{000}, p_{010}, p_{100}, p_{110}),$$
$$\sim \text{Dir}(\underline{\alpha}),$$

$$\underline{\alpha} = (\alpha_{000}, \alpha_{010}, \alpha_{100}, \alpha_{110}, \alpha_{000}, \alpha_{010}, \alpha_{100}, \alpha_{110})$$
$$= (1, 1, 1, 1, 1, 1, 1, 1)$$

Bayesian Analysis

If complete data

$$\underline{x} = (x_{000}, x_{010}, x_{100}, x_{110}, x_{000}, x_{010}, x_{100}, x_{110}),$$
$$\underline{y} = (y_{000}, y_{010}, y_{100}, y_{110}, y_{000}, y_{010}, y_{100}, y_{110}),$$

were available, then the likelihood is

$$\underline{x} \mid \underline{p} \sim \text{Mult}(x_{\square\square\square}, \underline{p}), \quad \underline{y} \mid \underline{p} \sim \text{Mult}(y_{\square\square\square}, \underline{p})$$

and the posterior distribution is

$$\underline{p} \mid \underline{x}, \underline{y} \sim \text{Dir}(\underline{\alpha} + \underline{x} + \underline{y})$$

Bayesian Analysis

Use Gibbs sampler, filling in missing data with predictive draws (data augmentation):

$$x_{s1d} \mid x_{s\Box d} \sim \text{Bin}(x_{s\Box d}, p_{s1d} / p_{s\Box d}), s, d = 0, 1$$

$$y_{st1} \mid y_{st\Box} \sim \text{Bin}(y_{st\Box}, p_{st1} / p_{st\Box}), s, t = 0, 1$$

$$x_{s0d} = x_{s\Box d} - x_{s1d}, \quad y_{st0} = y_{st\Box} - y_{st1}$$

$$\underline{p} \mid \underline{x}, \underline{y} \sim \text{Dir}(\underline{\alpha} + \underline{x} + \underline{y})$$

Iterative sampling converges to

$$\underline{p}, \underline{x}_{\text{miss}}, \underline{y}_{\text{miss}} \mid \underline{x}_{\text{obs}}, \underline{y}_{\text{obs}}$$

Objective Response Rates

| | Posterior | | |
|------------------|-----------|------|--------------|
| MRT ₁ | Mean | SD | 95% CI |
| – | 38.2% | 3.7% | 31.2, 45.5% |
| + | 53.9% | 2.0% | 50.0, 57.8% |
| Diff | -15.8% | 4.2% | -23.8, -7.5% |

| | Posterior | | |
|------------------|-----------|-------|--------------|
| MRT ₂ | Mean | SD | 95% CI |
| – | 43.9% | 15.3% | 15.3, 70.0% |
| + | 53.8% | 8.3% | 39.2, 69.8% |
| Diff | -9.9% | 4.2% | -53.2, 30.0% |



Review Issues

- Before applying IVD to clinical samples,
 - Finalize and lock down IVD, including cut-off.
 - Characterize analytical performance of IVD.
- Validate IVD on independent dataset.
- For targeted trials, IVD effectiveness cannot be evaluated.
- Include patients with missing test results in an *Intent to Diagnosis* analysis.

Review Issues

- **Bridging**

- Strongly prefer trial to external samples
- Need to assess concordance of MRT on LDT –'s as well LDT +'s.
- Missing samples could introduce bias
- Assess concordance near cut-off
- Use statistical methods to estimate drug efficacy in MRT defined subgroup, accounting for uncertainty due to discordance with LDT (if possible).

FDA Review Issues

- **Prospective Retrospective Analysis.**
 - Analytical performance of test is acceptable.
 - High ascertainment of test results
 - Trial not initially not a failure
 - Retrospective analysis plan specifically designed to assess efficacy in a defined subset.
 - Retrospective analysis not to be repeated.

FDA Guidances

- Drug-Dx Co-Development Concept Paper, 2005.
<http://www.fda.gov/cder/genomics/pharmacoconceptfn.pdf>
- Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests
<http://www.fda.gov/cdrh/osb/guidance/1620.html>
- FDA IVD MIA Draft Guidance
<http://www.fda.gov/cdrh/oivd/guidance/1610.html>
- FDA Draft Guidance for the Use of Bayesian Statistics in Medical Devices
www.fda.gov/cdrh/osb/guidance/1601.html

