Identifying a Subpopulation for a Tailored Therapy: Bridging Clinical Efficacy From a Laboratory-Developed Assay to a Validated In Vitro Diagnostic Test Kit

Published in: Statistics in Biopharmaceutical Research
February 2014, Vol. 6, No. 1, p.78-88

Jonathan Denne a, Gene Pennello b, Luping Zhao a, Shao-Chun Chang a & Sandra Althouse c

a Eli Lilly & Co., Indianapolis, IN, 46285  
b Diagnostics Devices Branch, Division of Biostatistics, OSB/CDRH/FDA  
c Work done while at Eli Lilly & Co.

Mid-West Biopharmaceutical Workshop  
May 2015
On July 6, 2012, FDA approved Erbitux® (cetuximab) for use in combination with FOLFIRI for first-line treatment of patients with \textit{K-ras} mutation-negative (wild-type), EGFR-expressing metastatic colorectal cancer (mCRC) as determined by FDA-approved tests for this use.

FDA approved concurrently the \textit{Therascreen KRAS RGQ PCR Kit} for identifying patients with \textit{K-ras} ‘wild-type’ (WT) mCRC.

The approval was based on retrospective analyses in the patient subsets according to \textit{K-ras} mutation status in tumor samples from patients enrolled in the CRYSTAL trial and in two supportive studies.

A key component was a study bridging the laboratory assay originally used in CRYSTAL to the \textit{Therascreen In-Vitro Diagnostic (IVD) kit}. 
Contents

♦ Background of cetuximab in mCRC
♦ Why a bridging study was needed
♦ Design of the Bridging study
♦ Results
♦ Summary
Background
Erbitux® in the treatment of Metastatic Colorectal Cancer (mCRC)

-US-Approved Indications in mCRC:
  - in combination with irinotecan in patients refractory to irinotecan (2004 - AA)

-The registration study, CRYSTAL, for Erbitux® as an initial treatment for mCRC was initiated in 2004
CRYSTAL Study

DESIGN*
- Phase 3, randomized, open-label, multicenter, controlled study
- Primary objective: PFS
- Key secondary objectives: OS, RR, safety, and QoL
- 1st-line mCRC (EGFR +ve)
- N=1217 randomized

RESULTS
- Significant PFS benefit (June 2006)
- No statistically significant survival benefit. (Dec 2007)

* Study Design did not include testing for any biomarkers apart from EGFR.

Abbreviations: OS = overall survival; PFS = progression-free survival; ORR=Overall Response Rate; QoL=Quality of Life
Emerging Data- Predictive Biomarker

♦ Potential signal for \textit{K-Ras} as a predictive biomarker emerging in 2007

♦ Voluntary Genomics Data Submissions (VGDS) to FDA of analyses by \textit{K-Ras} status in 5 Erbitux® mCRC studies (April 2008)

♦ \textit{K-Ras} mutation status added to Erbitux® label in Europe (July 2008)

♦ NCCN Guidelines recommend restricting use of Erbitux® to \textit{K-Ras} wild-type population (Oct 2008).

♦ FDA convened an ODAC on biomarkers, using \textit{K-Ras} as an example (Dec 2008), and subsequently (July 2009) issued class labeling recommending anti-EGFR mAb’s not be used in \textit{K-Ras} mutant tumors (safety language)
CRystAL Results by K-Ras Status

Following 2008 ODAC, K-Ras testing (increased sample ascertainment of 89%) was conducted, and an updated OS lock performed (May 2009).

<table>
<thead>
<tr>
<th></th>
<th>Erbitux®+ FOLFIRI (N=608)</th>
<th>FOLFIRI (N=609)</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>K-Ras Wild-Type Population† (N=676)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS</td>
<td>9.5 mo</td>
<td>8.1 mo</td>
<td>0.70</td>
<td>(0.57, 0.86)</td>
</tr>
<tr>
<td>OS*</td>
<td>23.5 mo</td>
<td>19.5 mo</td>
<td>0.80</td>
<td>(0.67, 0.94)</td>
</tr>
<tr>
<td><strong>K-Ras Mutant Population† (N=403)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS</td>
<td>7.5 mo</td>
<td>8.2 mo</td>
<td>1.13</td>
<td>(0.88, 1.46)</td>
</tr>
<tr>
<td>OS*</td>
<td>16.0 mo</td>
<td>16.7 mo</td>
<td>1.04</td>
<td>(0.84, 1.29)</td>
</tr>
</tbody>
</table>

*Post-hoc updated OS analysis, results based on an additional 162 events. †Based on 89% ascertainment rate for K-Ras testing. Abbreviations: HR = hazard ratio; OS = overall survival; PFS = progression-free survival.

OS and PFS benefit in K-Ras wild-type population, and no benefit in K-Ras mutant population.
Based on the Dec 2008 ODAC and subsequent meetings with FDA, a regulatory path forward emerged:

• Submission of retrospective analyses by K-ras mutation status
  • in the CRYSTAL study as pivotal and
  • in two supportive studies showing predictive nature of K-Ras
    – CA225-025 (Ph3 study of Erbitux® v BSC in previously treated mCRC) and
    – OPUS (Ph2 study of Erbitux® in combination with FOLFOX-4 1st line mCRC)

However, K-Ras testing in CRYSTAL had been done using a lab assay and a validated IVD kit would need to be developed for K-Ras testing and would be required for approval

→ a study would be needed to bridge between the lab assay and the IVD kit

First application of retrospective analysis in regulatory path for approval of a drug and companion diagnostic
Why was a Bridging study needed?
Why was a Bridging study needed?

CRYSTAL

Target population: K-Ras wild-type population

Original Diagnostic tool: lab assay (LNA assay)

Diagnostic tool: Therascreen K-Ras IVD kit
Why was a Bridging study needed?

CRYSTAL

Target population: K-Ras wild-type population

Original Diagnostic tool: lab assay (LNA-mediated qPCR clamping assay)

H owever <30% sample are available for retesting

Diagnostic tool: Therascreen K-Ras IVD kit

Bridge the CRYSTAL efficacy results from a K-Ras wt (LNA assay) population to a K-Ras wt (therascreen Kit) population.
### Initial Team Thoughts

#### LNA Assay

<table>
<thead>
<tr>
<th>K-RAS mutation status</th>
<th>Wild-Type</th>
<th>Mutant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-Type</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Mutant</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>Total</td>
<td>a+c</td>
<td>b+d</td>
<td>N</td>
</tr>
</tbody>
</table>

Agreement = \( \frac{(a+d)}{N} \)

Acceptable threshold?
Motivated by Pennello (2010)

We have:

CRYSTAL K-Ras wt (LNA) population
OS Hazard Ratio
= 0.80 (0.67 – 0.94)
Motivated by Pennello (2010)

We have:

CRYSTAL $K$-Ras wt (LNA) population
OS Hazard Ratio
$= 0.80 \ (0.67 – 0.94)$

We want:

Concordance between the two $K$-Ras tests

CRYSTAL $K$-Ras wt (TH) population
OS Hazard Ratio
$= ?? \ (?? – ??)$
Bridging Framework

Motivated by Pennello (2010)

We have:
- CRYSRAL K-Ras wt (LNA) population
  OS Hazard Ratio = 0.80 (0.67 – 0.94)

Concordance between the two K-Ras tests

We want:
- CRYSRAL K-Ras wt (TH) population
  OS Hazard Ratio = ?? (?? – ??)

Advantages
- Leads to an objective threshold for concordance
  - minimum concordance for which the estimated OS HR in the K-Ras wt (TH) population is statistically significant.
- An estimate (and CI) for OS HR in the K-Ras wt (TH) population is informative
Under certain assumptions, it can be shown that

- The OS HR in the K-Ras WT(TH) population depends on
  \[ p = \frac{a}{a+b} \] ← primary concordance measure

- The threshold for \( p \) was derived, giving \( H_0: p=0.80 \).

- As only WT(TH) will receive Erbitux®, choice of primary measure is consistent with clinical interest in the % WT(TH) patients who are MT(LNA).
Bootstrapping with Imputation

1. Create bootstrap dataset: sample pts with replacement from *K-Ras* evaluable population;
   - complete cases: patients with *K-Ras* mutation status by both tests
   - incomplete cases: patients with *K-Ras* mutation status (LNA) only

2. Fill-in each missing *K-Ras* test result for TH test by random selection among patients with observed TH test result and same LNA test result.

3. Compute the log hazard ratio based on filled-in bootstrap dataset.

4. Repeat the 3 steps of bootstrap, impute, and analyze 1000 times. Summarize the mean, SD, and 95% CI of the log hazard ratios.

CRYSTAL Study

| All Randomized (N=1217) | K-Ras (TH) & (LNA) (n=298) | K-Ras (LNA) only (n=765) | NA |

- All Randomized (N=1217)
- K-Ras (TH) & (LNA) (n=298)
- K-Ras (LNA) only (n=765)
- NA
To impute unknown $K$-$Ras$ (TH) based on $K$-$Ras$ (LNA) assuming Missing at Random (MAR), and then estimate OS hazard ratio in $K$-$Ras$ (TH) patients.

**CRYSTAL Study**

- **All Randomized** (N=1217)
- **$K$-$Ras$ (LNA)** (n=1063)
- **$K$-$Ras$ (TH)** (n=298)
  - Impute $K$-$Ras$ (TH) (n=765)

Simple imputation model: $K$-$Ras$ mutation status (LNA) only.

Imputation model (VO): $K$-$Ras$ mutation status (LNA) + pre-specified list of variables that might affect $K$-$Ras$ results + OS-time + OS-censoring-indicator + treatment

Full imputation model (VHI): same as VO model above but replacing OS time with Nelson-Aalen estimator of the cumulative hazard $H(T)$ and adding interactions with treatment
Representativeness of the Bridging Population
Recall that <30% samples were available for retesting using the *therascreen* Kit – a potentially non-random sample.

Validity of bridging rests on the samples available for retesting with the *therascreen* kit being representative of the whole study population.

- Specifically:
  - Identifying a list of variables that could affect the test result
  - Assessing any systematic differences between bridging sample and the non-bridging sample, especially in any variables associated with lower/higher levels of concordance between the 2 tests,
  - Evaluating impact of any observed differences on the bridging study results by inclusion in the multiple imputation model
List of the Variables

♦ Patient characteristics:
  • Gender, Race, Age, etc…

♦ Disease characteristics:
  • Months from first histological diagnosis to randomization, Primary diagnosis (rectum only vs. colon), etc…

♦ Immediate characteristics of tumor sample:
  • Tumor type, Tumor content in sample, etc…

♦ Handling and processing factors:
  • Age of sample at testing, Sampling method, Enrollment site, Region, etc…
Results of the Bridging Study
Concordance Between the QIAGEN *therascreen* Kit and LNA-Mediated qPCR Clamping Assay

<table>
<thead>
<tr>
<th><em>K-Ras</em> mutation status</th>
<th>LNA Assay</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wild-Type</td>
<td>Mutant</td>
</tr>
<tr>
<td><strong>Wild-Type</strong></td>
<td>141</td>
<td>7</td>
</tr>
<tr>
<td><strong>Mutant</strong></td>
<td>30</td>
<td>120</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>171</td>
<td>127</td>
</tr>
</tbody>
</table>

Primary concordance measure  \( p = \frac{141}{148} = 0.953 \) (95% CI: 0.91-0.98)

Note: Overall Disagreement (30+7=12.4%) mostly arising from the TH kit identifying samples as MT, which the LNA assay did not.
# Estimated vs. Observed OS Results in CRYSTAL  K-Ras Wild-Type Population

<table>
<thead>
<tr>
<th>Method</th>
<th>OS Hazard Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>By LNA-mediated qPCR clamping assay</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRYSTAL (observed)</td>
<td>0.796</td>
<td>0.670, 0.946</td>
</tr>
<tr>
<td><strong>By QIAGEN therascreen Kit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytic method</td>
<td>0.806</td>
<td>0.678, 0.958</td>
</tr>
<tr>
<td>Bootstrapping with imputation</td>
<td>0.800</td>
<td>0.668, 0.958</td>
</tr>
<tr>
<td>Multiple imputation (simple model)</td>
<td>0.807</td>
<td>0.659, 0.987</td>
</tr>
<tr>
<td>Multiple imputation (VO model)</td>
<td>0.805</td>
<td>0.658, 0.984</td>
</tr>
<tr>
<td>Multiple imputation (VHI model)</td>
<td>0.786</td>
<td>0.638, 0.968</td>
</tr>
</tbody>
</table>
Sensitivity analyses

“For each value of $k$, add the following step to the bootstrap analysis:
• Randomly select "$k" number of patients and change the $K$-$Ras$ status of these patients

• $k=18$ is the highest value of $k$ for which the upper 95%CI for the estimated OS HR in the $K$-$Ras$ wt (TH) population was <1, supporting robustness of results.

“Tipping point analysis”* - systematically decrease the concordance between to examine how many addition samples would need to be discordant to lose statistical significance for the estimated OS hazard ratio in the $K$-$Ras$ wt (TH) patients.

<table>
<thead>
<tr>
<th>LNA-Mediated qPCR Clamping Assay</th>
<th>K-RAS mutation status</th>
<th>Wild-Type</th>
<th>Mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIAGEN therascreen Kit</td>
<td>Wild-Type</td>
<td>$k=141$</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Mutant</td>
<td>30</td>
<td>120</td>
</tr>
</tbody>
</table>

*Yan et al. (2009)
Rigorous statistical design and analysis of a bridging study enabled the efficacy from a pivotal study in a population identified by a lab assay to be bridged to a population identified by the proposed companion diagnostic.

In particular,

- Statistical models to enable the estimation of the efficacy in the intended indicated population (i.e. based on the IVD kit)
- Leading to an objective definition of the minimally required concordance needing to be demonstrated in the bridging study


