Performance Variability of BCR-ABL Monitoring Tests: Impact of PCR Platform Standardization on Health Care Systems

Poluru L. Reddy[1], Paul Choppa[2], Mehdi R. Jamehdor[3], Mike M. Moradian[3], Nicholas T. Potter[4], Michael M. Quigley[5], Christopher D. Watt[6], Scott A. Campbell[7], Wendy W. Wong[8], Holger A. Höfling[9], Brian M. Manning[10], Andre A. Mignault[10], Diego F. Ossa[9], Andrew M. Stein[10], Peter Stiegler[11], Sha-Sha Wang[10], Feng Yang[10], Meir Wetzler[12]

[1]University of Chicago Medical Center, Chicago, Illinois, USA; [2]Clarient, a GE Healthcare Company, Aliso Viejo, California, USA; [3]Kaiser Permanente Southern California, Los Angeles, California, USA; [4]Molecular Pathology Laboratory, Inc., Maryville, Tennessee, USA; [5]Scripps Green Hospital, La Jolla, California, USA; [6]Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, USA; [7]Cepheid, Rolling Meadows, Illinois, USA; [8]Cepheid, Sunnyvale, California, USA; [9]Novartis Pharma AG, Basel, Switzerland; [10]Novartis Institutes for BioMedical Research, Cambridge, Massachusetts, USA; [11]Duke Clinical Research Institute, Durham, North Carolina, USA; 12Roswell Park Cancer Institute, Buffalo, New York, USA

INTRODUCTION

- Molecular monitoring of BCR-ABL transcript levels by real-time quantitative polymerase chain reaction (RQ-PCR) is the most sensitive method for measuring minimal residual disease burden for patients with chronic myeloid leukemia (CML)[1]
- Molecular response is an important aspect of successful CML therapy[1] and is included in current guidelines and recommendations published by the National Comprehensive Cancer Network (NCCN) and the European LeukemiaNet (ELN)
- BCR-ABL tests are often independently designed and developed at individual laboratories potentially introducing sources of variation[2,3]
- An International Scale (IS) has been established to improve the reproducibility and accuracy of molecular monitoring[4-6]; however, not all laboratories have established conversion factors[7]

STUDY OBJECTIVES

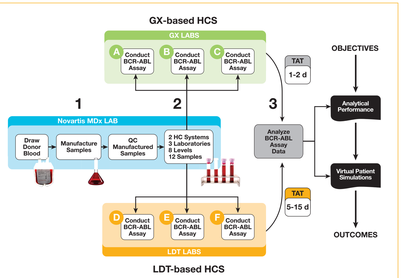
REVEAL (Reproducibility and Variability Evaluation of Assays in Leukemia)

- Study objectives:
  - Explore intralaboratory and interlaboratory variability among BCR-ABL monitoring tests by evaluating the impact of PCR platform standardization using the GeneXpert® BCR-ABL Assay (GX) (Cepheid, Sunnyvale, California) standardized to the IS and laboratory-developed tests (LDTs) not standardized to the IS
  - Compare analytical performance of BCR-ABL monitoring tests in 2 health care systems (HCSs)
    - A GX-based health care system that relies on 3 sites using an automated and IS-standardized BCR-ABL test
    - An LDT-based health care system relying on 3 independently developed BCR-ABL tests
  - Compare the budget impact between GX-based HCS and LDT-based HCS

STUDY OVERVIEW

1. CML patient analogue samples were prepared corresponding to targeted BCR-ABL^{IS} ratios ranging from ≈10% to ≈0.01%
2. Blinded samples were sent to 6 laboratories
   - 4 rounds of testing were performed in 2 phases
   - 3 laboratories used the GX (GX-based HCS)
   - 3 laboratories used their own internally developed and validated BCR-ABL quantification assays (LDT-based HCS)
3. Results were unblinded to compare the analytical performance of the individual laboratories and the GX- and LDT-based HCSs

Figure 1. Study Overview



Participating Laboratories

- Clarient, a GE Healthcare Company
- Hospital of the University of Pennsylvania Molecular Pathology Laboratory
- Molecular Pathology Laboratory Network, Inc.
- Medical Genetics Laboratories, Kaiser Permanente Southern California
- Scripps Clinic Medical Laboratories
- The University of Chicago Medical Center, Molecular Diagnostics Laboratory

METHODS

Phase I

- Phase I included round A on November 2, 2010, and round B on November 9, 2010 (Table 1)
- Target concentrations of K562 cells were selected based on an empirically determined regression of the relationship between cell concentration and BCR-ABL^{IS} % ratio analytical response as measured on the GX platform
- Targeted BCR-ABL^{IS} levels in phase I corresponded to approximately 10.0%, 1.0%, 0.1%, and 0

Phase II

- Phase II included round C on February 8, 2011, and round D on February 15, 2011 (Table 1)
- Unlike phase I, analytical response near major molecular response (MMR) was benchmarked to a specific cell ratio (9.0 K562 cells/million white blood cells [WBC]) and was independently verified at 2 separate reference laboratories
  - Targeted MMR sample (BCR-ABL^{IS} = 0.1%)
    - Tested by the Department of Molecular Pathology, Centre for Cancer Biology, SA Pathology, Adelaide, Australia → BCR-ABL^{IS} = 0.12%
    - Tested by the Department of Pathology, Molecular Diagnostics Laboratory, University of Toronto Health Network, Toronto, Ontario, Canada → BCR-ABL^{IS} = 0.09%
- Samples were then prepared based on cell ratios at levels approximately one-half log below those targeted in phase I corresponding to approximately 3.16%, 0.32%, 0.03%, and 0.01%

Table 1. Discrete Target Analyte Levels of K562 in Normal Whole Blood Assayed Across GX and LDT Health Care Systems

Level	Approx. Clinical Level	IS Log Reduction	Target BCR-ABL^{IS} Ratio	WBC Per mL whole blood, n	K562 Cells Per mL blood, n	Cell Ratio (K562 Cells/Million WBC, n)	Phase	Round[a]
High	CHR	1.0	10.00%	6.7	6014.0	897.6	I	A
	≥ CCyR	1.5	3.16%	8.5	3315.0	390.0	II	C
Medium	CCyR	2.0	1.00%	6.8	601.4	88.4	I	A
	≥ MMR	2.5	0.32%	5.6	218.4	39.0	II	C
Low	MMR	3.0	0.10%	4.3	80.0	18.6[b]	I	B
	MR$^{3.5}$	3.5	0.03%	6.0	17.9	3.9	II	D
Very low	MR$^{4.0}$	4.0	0.01%	4.6	7.4	1.2	II	B
Negative	CMR	-	0%	3.7	0.0	0.0	I	D

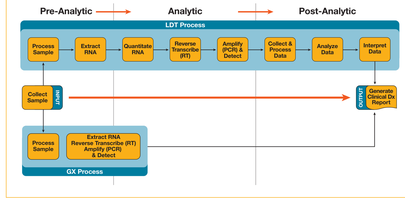
CCyR, complete cytogenetic response; **CHR**, complete hematologic response; **CMR**, complete molecular response; **IS**, international scale; **MMR**, major molecular response; **MR^{x}**, BCR-ABL^{IS} ≤ 0.01%; **MR$^{4.5}$**, BCR-ABL^{IS} ≤ 0.0032%; **WBC**, white blood cells.

[a]Round A: November 2, 2010; Round B: 9 November, 2010; Round C: February 8, 2011; Round D: February 15, 2011. [b]A technical error in round B resulted in 33% excess concentration of K562 for the "low" analyte level, targeting MMR at BCR-ABL^{IS} = 0.1%. Coupled with a low WBC count in the donor blood used to manufacture samples for this analyte level, the excess K562 induced an approximately 2-fold increase in observed BCR-ABL^{IS} % ratio response on the GX platform, above the nominal target of BCR-ABL^{IS} = 0.1%. The actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS).

BCR-ABL Monitoring Assay Process

- GX-based HCS
  - Samples were tested according to the manufacturer's instructions
  - Results reported on the IS were derived by applying the manufacturer's validated conversion factor established for its CE IVD Xpert® BCR-ABL Monitor test (Cepheid, Sunnyvale, California)
- LDT-based HCS
  - Samples were tested according to individually developed BCR-ABL quantification assays per the testing sites' standard operating procedures
  - Laboratories in the LDT-based HCS had not undertaken the process of obtaining an IS conversion factor and consequently results are reported exactly as provided by those sites with the raw data transformed to % BCR-ABL/ABL ratios to facilitate data comparison
- In both HCSs, samples underwent all preanalytic, analytic, and postanalytic workflow steps (Fig 2)
- Qualitative results were reported for every sample and data were assigned into one of the following four categories:
  1. Positive (POS; LDT and GX): analytical result confirming presence of BCR-ABL transcript in a sample at levels that permitted quantitation expressed as a ratio (BCR-ABL/ABL)
  2. Nonquantifiable positive (NQP; LDT only): analytical result confirming presence of BCR-ABL transcript in a sample at levels that could not be quantitated as a ratio (BCR-ABL/ABL) (ie, below limit of quantitation for the assay)
  3. Negative (NEG; LDT and GX): analytical result that did not detect the presence of BCR-ABL transcript in a sample in the context of a valid ABL signal (ie, below limit of detection for the assay)
  4. Failure (FAIL; LDT and GX): permanent and unrecoverable loss of clinical diagnostic data due to failure(s) during execution of preanalytical, analytical, and/or postanalytical sample processing

Figure 2. BCR-ABL Monitoring Assay Process



Statistical Analyses

- The Wilcoxon test was used to determine whether the distribution of measurements of 2 laboratories was identical
- The Levene test (with median) was used to determine whether the variances between 2 laboratories was identical
- Fisher exact test was used to test if distribution of qualitative results were identical for 2 laboratories
- A significance level $\alpha = 0.0018 = 0.05/(9 * 3)$ was used, given 9 pairwise laboratory comparisons, with 3 tests each, per the Bonferroni correction method to adjust for multiplicity

Table 2. Summary of Results for Each Laboratory by Target BCR-ABL Levels

HCS	Lab	Target BCR-ABL^{IS} Ratio: 10.00%	3.16%	1.00%	0.32%	0.10%*	0.03%	0.01%	Negative Control
		Median of POS Results (Range)[b]							
GX (IS)	Lab A	9.6 (3.8-11)	3.3 (0.17-4.7)	0.82 (0.47-1.2)	0.36 (0.25-0.85)	0.23 (0.13-0.41)	0.02 (0.0004-0.05) 2 NEG	0.006 (0.001-0.03) 8 NEG	12 NEG
	Lab B	9.9 (7.0-14)	2.3 (0.12-3.7)	0.99 (0.75-1.6)	0.34 (0.11-0.56)	0.19 (0.05-0.30)	0.02 (0.0003-0.08) 3 NEG	0.003 (0.001-0.004) 9 NEG	11 NEG 1 FAIL
	Lab C	10 (8.5-13)	3.7 (2.4-6.1)	1.1 (0.61-1.6)	0.42 (0.04-0.80)	0.19 (0.09-0.42)	0.03 (0.003-0.14) 1 NEG	0.006 (0.002-0.01) 8 NEG	12 NEG
LDT	Lab D	20 (16-27)	8.6 (4.8-10)	1.8 (1.3-2.5)	1.15 (0.84-1.6)	0.51 (0.22-0.77)	0.05 (0.04-0.11) 1 NQP	0.02 (0.02-0.02) 11 NQP	11 NEG 1 NQP
	Lab E	2.8 (1.6-3.9)	1.7 (0.22-2.5)	0.23 (0.21-0.30)	0.18 (0.06-0.31) 1 NEG	0.05 (0.01-0.13)	0.008 (0.002-0.01) 1 NEG 5 NQP	0.01 (0.01-0.01) 7 NEG 4 NQP	12 NEG
	Lab F	29 (21-35)	9.40 (8.2-12)	2.1 (1.2-4.7)	0.90 (0.02-1.6) 1 FAIL	1.1 (0.65-2.1) 1 FAIL	0.12 (0.11-0.17) 8 FAIL	0.01 (0.003-0.03) 4 FAIL	0.009 (0.006-0.03) 7 NEG

FAIL, failure; **GX**, automated GeneXpert System; **HCS**, health care system; **IS**, International Scale; **LDT**, laboratory-developed tests; **NEG**, negative; **NQP**, nonquantifiable positive; **POS**, positive.

[a]During panel manufacturing, the actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS). [b]Median and range of positive results; 12 replicate measurements per level per laboratory.

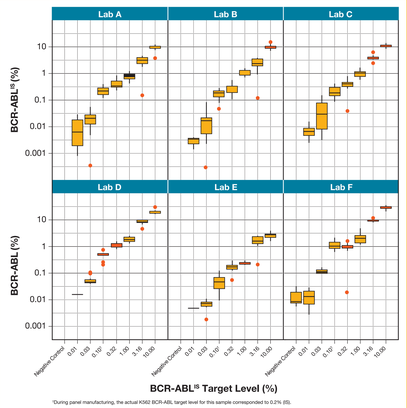
Table 3. Statistical Comparisons of Results of All Laboratories

	Wilcoxon	Levene	Fisher
GX Laboratories			
Lab A vs Lab B	0.043	0.69	1
Lab A vs Lab C	0.16	0.3	1
Lab B vs Lab C	0.015	0.72	0.98
LDT Laboratories			
Lab D vs Lab E	1.7 x 10^{-19a}	0.0072	0.0077
Lab D vs Lab F	5 x 10^{-7}	0.21	0.0002[b]
Lab E vs Lab F	7.1 x 10^{-19a}	0.00082[a]	0.0002[b]
GX^{IS} vs LDT Laboratories			
GX Labs vs Lab D	0.00077*	0.016	1.8 x 10^{-9a}
GX Labs vs Lab E	4.3 x 10^{-23a}	0.011	0.0002[b]
GX Labs vs Lab F	5.5 x 10^{-11a}	0.22	2.9 x 10^{-5a}

[a]Significant difference [$\alpha = 0.0018 = 0.05/(9 * 3)$] [b]*P* values for GX laboratories combined vs LDT laboratories using all rounds.

- Ratio of median measurements for each target level was < 2 for the GX sites and a range of ≈10 to 22 for the LDT sites
- None of the 3 GX laboratory comparisons were significant for Wilcoxon, Levene, or Fisher exact tests on level $\alpha = 0.0018$
- Mean measurement levels for all LDT laboratories differed significantly for all levels (tested by Wilcoxon)

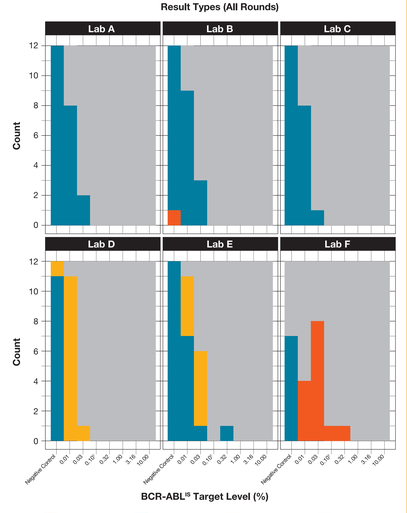
Figure 3. Results for Each BCR-ABL Target Level by Laboratory



*During panel manufacturing, the actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS).

- Overall, a high level of agreement was observed among laboratories using the GX system standardized to the IS at each of the analyte levels tested

Figure 4. Qualitative Sample Results by Laboratory



*During panel manufacturing, the actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS).

- A subset of matched replicate samples with low BCR-ABL levels were reported as positive but not quantifiable by LDT laboratories (21 of 48 replicates)
  - Low BCR-ABL levels were positive but not quantifiable in laboratories D and E
  - Laboratory F reported half of the replicates for low-level samples as failures (12/24) because of a malfunctioning instrument and all others as positives (12/24)
- A subset of samples with low levels of BCR-ABL (below MMR) was reported as negative by the GX system (31/72 replicates)

Table 4. Summary of Results by HCS

	Median of POS Results (Range)[a]							
	10.00%	3.16%	1.00%	0.32%	0.10%[b]	0.032%	0.01%	Negative control
GX-based HCS	9.9 (3.8-14)	3.5 (0.12-6.1)	0.94 (0.47-1.6)	0.35 (0.04-0.85)	0.21 (0.05-0.42)	0.02 (0.0003-0.14) 6 NEG	0.004 (0.001-0.03) 25 NEG	35 NEG 1 FAIL
LDT-based HCS	20 (1.6-35)	8.3 (0.22-12)	1.5 (0.21-4.7)	0.90 (0.02-1.6) 1 FAIL 1 NEG	0.5 (0.01-2.1) 1 FAIL	0.05 (0.002-0.17) 1 NEG 6 NQP 8 FAIL	0.01 (0.003-0.03) 7 NEG 15 NQP 4 FAIL	0 .009 (0.006-0.03) 30 NEG 1 NQP

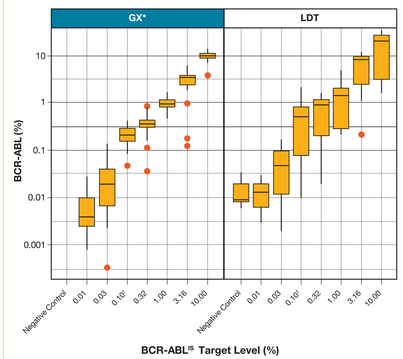
FAIL, failure; **HCS**, health care system; **GX**, automated GeneXpert System; **LDT**, laboratory-developed tests; **NEG**, negative; **NQP**, nonquantifiable positive; **POS**, positive.

[a]Median and range of positive results; 36 replicate measurements per level per health care system.

[b]During panel manufacturing, the actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS).

- A 10-fold variation existed among samples in the LDT-based HCS, whereas fewer differences were observed for the IS standardized and automated GX-based HCS

Figure 5. Results for Each BCR-ABL Target Level by HCS



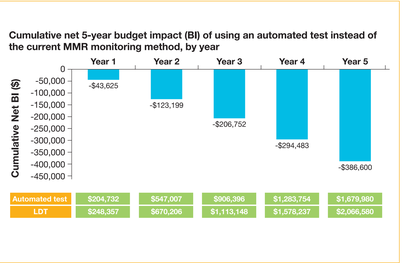
*Results reported per the IS.
*During panel manufacturing, the actual K562 BCR-ABL target level for this sample corresponded to 0.2% (IS).

- Considerable overlap was observed among different BCR-ABL levels in the samples tested within an LDT-based HCS
- Overlap was generally not observed for the GX-based HCS (with the exception of the BCR-ABL^{IS} = 0.01% and 0.03% samples)

Economic Impact

- A model to estimate the budget impact (BI) with an automated BCR-ABL monitoring test compared with an LDT was developed for newly diagnosed patients with CML in the United States
- The full details and assumptions of this model have been previously reported[7]
  - The model used best available data
  - Epidemiology data were combined with workflow costs based on a survey of US laboratories
  - A testing algorithm based on NCCN guidelines was used to capture a number of different monitoring tests for CML, including determination of MMR

Figure 6. Cost Savings Over 5 Years With an Automated BCR-ABL Monitoring Test vs an LDT



- For 100 patients whose BCR-ABL monitoring is being followed according to NCCN guidelines, an overall savings of $43,625 could be achieved in 1 year and $386,600 over 5 years when implementing an automated test

CONCLUSIONS

- Statistical evaluation of the GX results demonstrated no significant differences across the 3 testing sites
- Results significantly varied between the 3 laboratories each using a unique LDT. Although within a laboratory most LDT results were reproducible, some LDTs consistently under- or overreported BCR-ABL transcript levels with respect to the IS
- Variability in the LDT-based HCS was considerably higher than in the GX-based HCS at levels near or above MMR
- Automated BCR-ABL testing not only reduces interlaboratory variability, but represents an estimated $387,000 cost-saving alternative versus LDTs over a 5-year period
- These results underscore the need for PCR standardization and suggest that automated platform standardization and adoption of the IS can improve BCR-ABL molecular monitoring for patients with CML
- Interlaboratory variation and differences in BCR-ABL ratios may impact clinical disease management in patients with CML
  - An erroneous assignment of MMR may prevent guideline-recommended follow-up to investigate potential problems, such as patient noncompliance or possible early emergence of TKI-resistant disease
  - An erroneous report indicating MMR has not been achieved could trigger unnecessary follow-up, such as additional bone marrow aspiration and biopsy for cytogenetics or repeated PCR testing

REFERENCES

1. Baccarani M, et al. *J Clin Oncol.* 2009;27(35):6041-6051.
2. Branford S, et al. *Blood.* 2008;112(8):3330-3338.
3. Zhang T, et al. *J Mol Diagn.* 2007;9(4):421-430.
4. Branford S, et al. *Leukemia.* 2006;20(11):1925-1930.
5. Müller MC, et al. *Leukemia.* 2009;23(11):1957-1963.
6. Hughes, T et al. *Blood.* 2006;108(1):28-37.
7. Ratcliffe A, et al. *Value Health.* 2011;14(7):A247.

Poster Presentation at the Association for Molecular Pathology 2011 Annual Meeting, November 17-19, 2011, Grapevine, TX.