Evaluation of MMR Concordance Based on Either International Scale or National Comprehensive Cancer Network Criteria Using Reconstructed "Virtual" CML Patient Profiles From the REVEAL BCR-ABL Methods Comparison Study

Poluru L. Reddy, PhD, DABCC[1]; Paul Choppa[2]; Mike M. Moradian, PhD[3]; Nicholas T. Potter, PhD, FACMG[4]; Michael M. Quigley, MD[5]; Christopher D. Watt, MD, PhD[6]; Holger Höfling, PhD[7]; Brian Manning, PharmD[8]; Andre Mignault, MS[8]; Brian Mullaney, MD, PhD[8]; Diego Ossa, MD, MSc[9]; Andrew Stein, PhD[10]; Sha-Sha Wang, PhD[8]; Feng Yang, PhD[8]; Meir Wetzler, MD[11]

[1]Department of Molecular Diagnostics, University of Chicago Medical Center, Chicago, IL; [2]Clarient, a GE Healthcare Company, Aliso Viejo, CA; [3]Medical Genetics Laboratory, Kaiser Permanente Southern California, Los Angeles, CA; [4]Molecular Pathology Laboratory Network, Inc., Maryville, TN; [5]Pathology Department, Scripps Green Hospital, La Jolla, CA; [6]Hospital of the University of Pennsylvania, Philadelphia, PA; [7]Molecular Diagnostics, Novartis Pharma AG, Basel, Switzerland; [8]Novartis Molecular Diagnostics, Cambridge, MA; [9]Novartis Molecular Diagnostics, Basel, Switzerland; [10]Modeling & Simulation, Novartis Institutes for BioMedical Research, Cambridge, MA; [11]Dept. of Medicine, Leukemia Section, Roswell Park Cancer Institute, Buffalo, NY

BACKGROUND

- 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)
- Achieving major molecular response (MMR) is an important milestone in CML therapy
- MMR is defined as at least a 3-log reduction in BCR-ABL transcript levels (normalized to BCR endogenous control) from a laboratory-specific standardized baseline (BL), calculated as the median percent ratio among 30 newly diagnosed patients not yet receiving imatinib in the International Randomized Study of Interferon Versus STI571 (IRIS) trial[1]
- Standardization has been achieved through the development of an International Scale (IS), which defines MMR as BCR-ABL^{IS} = 0.1%[2-4]
- In contrast, the National Comprehensive Cancer Network (NCCN) defines MMR as a 3-log reduction in BCR-ABL transcript levels but is indefinite on the definition of BL[5]

STUDY OBJECTIVE

REVEAL (Reproducibility and Variability Evaluation of Assays in Leukemia)

- Explore intra-laboratory and inter-laboratory variability among BCR-ABL monitoring tests
- Evaluate the impact of PCR platform standardization using the GeneXpert® BCR-ABL Assay (GX) (Cepheid, Sunnyvale, California) standardized to the IS relative to laboratory-developed tests (LDTs) not standardized to the IS
- Compare analytical performance of BCR-ABL monitoring tests in 2 health care systems (HCS)
  - A GX-based HCS that relies on 3 sites using an automated and IS-standardized BCR-ABL test
  - An LDT-based HCS relying on 3 independently developed BCR-ABL tests
- The pairwise concordance of MMR determination was examined within and between 3 laboratories using the IS-standardized GX and 3 laboratories using LDTs

STUDY OVERVIEW

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

Concordance Determination

- 1000 virtual patients (VPs) were emulated based on data from the REVEAL BCR-ABL Methods Comparison Study
- VP emulations were guided by actual patient outcomes in landmark analyses of 7 treatment responses by Hughes et al.[6] For the analyses, patients were categorized as follows based on the achievement of molecular responses at each time point: BCR-ABL^{IS} ≤ 0.1%; BCR-ABL^{IS} > 0.1% to ≤ 1%; BCR-ABL^{IS} > 1% to ≤ 10%; BCR-ABL^{IS} ≥ 0.1%; MMR vs no MMR; and MMR vs BCR-ABL^{IS} > 0.1% to ≤ 10%
- Treatment response profiles over an 18-month time horizon were modeled by assigning 1 of the 7 BCR-ABL levels sampled in the REVEAL study to each of 4 virtual time points (eg, 3, 6, 12, and 18 months)
- BL levels were defined for each VP by drawing randomly from 2 different underlying distributions:

  Method 1: Uniform quartiles representing pretreatment BCR-ABL ratios between 50% and 150%, centered on 100% as defined by IS

  Method 2: 0.25-log increments observed for BL levels among 94 patients in the IRIS trial
- 6000 VP transcript profiles (VTPs) were then reconstructed using data from each of the 6 laboratories for all 1000 VPs
- The final 18-month time point in each VTP provided the BCR-ABL level against which the IS or NCCN objective criterion was applied to make MMR determinations
- MMR concordance was evaluated by inspecting all possible inter-laboratory pairwise comparisons among the 1000 VPs

Accuracy Determination

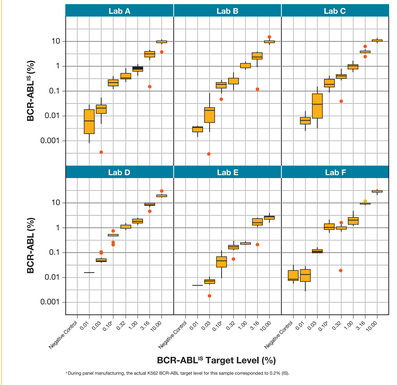
- Accuracy was evaluated by inspecting overall percent agreement between the rate of MMR determined by each laboratory and the rate of MMR estimated by the model
- The final 18-month time point in each VTP provided the BCR-ABL level against which the NCCN objective criterion was applied to make MMR determinations
  - MMR was achieved for a given VP if the 18-month value was at least 1000-fold lower than the BL diagnostic value modeled on IRIS BL data
- A nominal expected MMR status for a given VP was determined as a property of the discrete analyte level from REVEAL that was sampled at that 18-month time point, where only discrete analyte levels with central tendencies of BCR-ABL^{IS} < 0.1% were assigned an expected MMR status of "yes"
- Positive percent agreement (PPA) was calculated as the proportion of all nominally positive results that were correctly reported as positive by the laboratory
- Negative percent agreement (NPA) was calculated as the proportion of all nominally negative results that were correctly reported as negative by the laboratory
- Overall percent agreement (OPA) was calculated as the proportion of all VPs that were correctly reported as positive or negative by the laboratory

RESULTS

Analytical Results

- Ratio of median measurements for each target level was < 2 for the GX sites and a range of ≈10 to 22 for the LDT laboratories
- 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)

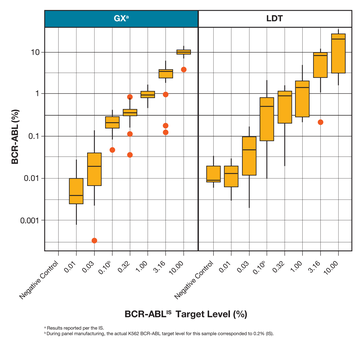
Figure 1. Results for Each BCR-ABL Target Level by Laboratory



*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 2. 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)

Simulation Results
Method 1 (Quartiles Around 100% IS)

Figure 3. Total % Pairwise Concordance: MMR Achieved at 18 Months per NCCN Objective Criterion

		A	B	C	D	E	F
GX Labs (NCCN)	A	.					
	B	96%	.				
	C	88%	87%	.			
LDT Labs (NCCN)	D	93%	90%	85%	.		
	E	85%	86%	76%	78%	.	
	F	52%	51%	59%	59%	37%	.

GX Labs (NCCN) | LDT Labs (NCCN)

- Pairwise concordance *in MMR as determined by NCCN criterion* among all 6 laboratories is shown in Figure 3
- MMR determinations among the 3 GX laboratories were concordant in 87% to 96% of VPs
- In contrast, MMR determinations among the LDT laboratories were concordant in 37% to 78% of VPs
- MMR determinations were concordant in 51% to 93% of VPs when compared between GX laboratories and LDT laboratories

Figure 4. Total % Pairwise Concordance: MMR Achieved at 18 Months per Mixed IS/NCCN Objective Criteria

		A	B	C	D	E	F
GX Labs (IS)	A	.					
	B	98%	.				
	C	93%	91%	.			
LDT Labs (NCCN)	D	93%	91%	88%	.		
	E	85%	86%	80%	78%	.	
	F	52%	50%	54%	59%	37%	.

GX Labs (IS) | LDT Labs (NCCN)

- When *MMR determination based on IS criterion for GX* and NCCN for LDTs was considered, MMR concordance improved to 91% to 98% among the GX laboratories in contrast to 50% to 93% concordance observed between the GX and LDT laboratories (Figure 4)
- It is noteworthy that Lab D results more closely approximated the IS than results from the other LDT laboratories examined in the REVEAL study
- Although Lab D does not report results per the IS, it does report results relative to a median diagnostic BL, similar to the approach used in Hughes et al[1]
- An LDT-based HCS without any attempted IS standardization resulted in MMR concordance of only 37% between Lab E and Lab F (Lab D removed from analysis)

Method 2 (IRIS)

Figure 5. Distribution of BCR-ABL BL levels from 94 patients with BL and 18-month assessments in IRIS

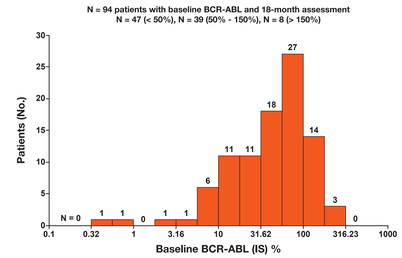


Figure 6. Total % Pairwise Concordance: MMR Achieved at 18 Months per NCCN Objective Criterion

		A	B	C	D	E	F
GX Labs (NCCN)	A	.					
	B	85%	.				
	C	80%	79%	.			
LDT Labs (NCCN)	D	79%	74%	79%	.		
	E	80%	81%	72%	61%	.	
	F	56%	53%	63%	76%	37%	.

GX Labs (NCCN) | LDT Labs (NCCN)

- *MMR as determined by NCCN criterion* among the 3 GX laboratories were concordant in 79% to 85% of VPs
- In contrast, MMR determinations among the LDT laboratories were concordant in 37% to 76% of VPs
- MMR determinations were concordant in 53% to 81% of VPs when compared between GX laboratories and LDT laboratories

Figure 7. Total % Pairwise Concordance: MMR Achieved at 18 Months per Mixed IS/NCCN Objective Criteria

		A	B	C	D	E	F
GX Labs (IS)	A	.					
	B	97%	.				
	C	93%	91%	.			
LDT Labs (NCCN)	D	68%	65%	67%	.		
	E	89%	87%	84%	61%	.	
	F	43%	40%	46%	76%	37%	.

GX Labs (IS) | LDT Labs (NCCN)

- When *MMR determination based on IS criterion for GX* and NCCN criterion for LDT laboratories was considered, MMR concordance improved to 91% to 97% among the GX laboratories in contrast to 40% to 89% concordance observed between the GX and LDT laboratories

Accuracy Results

Figure 8. Accuracy of MMR Determination

		Clinical Truth (per Model) MMR+	MMR–
GX Lab A Result[a]	MMR+	648	0
	MMR–	0	352

PPA = 100%; NPA = 100%; OPA = 100%

		Clinical Truth (per Model) MMR+	MMR–
GX Lab B Result[a]	MMR+	648	29
	MMR–	0	323

PPA = 100%; NPA = 91.8%; OPA = 97.1%

		Clinical Truth (per Model) MMR+	MMR–
GX Lab C Result[a]	MMR+	595	15
	MMR–	53	337

PPA = 91.8%; NPA = 95.7%; OPA = 93.2%

		Clinical Truth (per Model) MMR+	MMR–
LDT Lab D Result[a]	MMR+	324	0
	MMR–	324	352

PPA = 50.0%; NPA = 100%; OPA = 67.6%

		Clinical Truth (per Model) MMR+	MMR–
LDT Lab E Result[a]	MMR+	623	90
	MMR–	25	262

PPA = 96.1%; NPA = 74.4%; OPA = 88.5%

		Clinical Truth (per Model) MMR+	MMR–
LDT Lab F Result[a]	MMR+	81	567
	MMR–	0	352

PPA = 100%; NPA = 38.3%; OPA = 43.3%

[a] MMR determination based on IS.
[b] MMR determination based on NCCN criterion.

- For the GX system, the PPA ranged from 91.8% to 100%, the NPA ranged from 91.8% to 100%, and the OPA ranged from 93.2% to 100%
- For the LDT system, the PPA ranged from 50.0% to 100%, the NPA ranged from 38.3% to 100%, and the OPA ranged from 43.3% to 88.5%

CONCLUSIONS

- These results illustrate that the NCCN criterion for MMR determination is not adequate for inter-laboratory comparisons of BCR-ABL transcript levels near the clinically important level of MMR
- In contrast, standardization to the IS improves inter-laboratory concordance in MMR determination
- These results highlight the discrepancies that may result when comparing molecular responses between laboratories not standardized to the IS
- When using this IRIS-based model for assessing clinical truth, the use of an automated platform standardized to the IS improved the overall accuracy of determining MMR vs non-IS LDT methods
- As attainment of MMR is a critical milestone of CML therapy; errors in MMR determination may have an adverse impact on CML disease management

REFERENCES

1. Hughes TP, et al. *N Engl J Med.* 2003;349(15):1423-1432.
2. Branford S, et al. *Leukemia.* 2006;20(11):1925-1930.
3. Müller MC, et al. *Leukemia.* 2009;23(11):1957-1963.
4. Hughes T, et al. *Blood.* 2006;108(1):28-37.
5. O'Brien S, et al. *J Natl Compr Canc Netw.* 2011;9(suppl 2):S1-S25.
6. Hughes TP, et al. *Blood.* 2010;116(19):3758-3765.

QR code

Poster Presentation at the 53rd American Society of Hematology Annual Meeting and Exposition, December 10-13, 2011, San Diego, California