

blood

2008 112: 3330-3338
Prepublished online August 6, 2008;
doi:10.1182/blood-2008-04-150680

Desirable performance characteristics for *BCR-ABL* measurement on an international reporting scale to allow consistent interpretation of individual patient response and comparison of response rates between clinical trials

Susan Branford, Linda Fletcher, Nicholas C. P. Cross, Martin C. Müller, Andreas Hochhaus, Dong-Wook Kim, Jerald P. Radich, Giuseppe Saglio, Fabrizio Pane, Suzanne Kamel-Reid, Y. Lynn Wang, Richard D. Press, Kevin Lynch, Zbigniew Rudzki, John M. Goldman and Timothy Hughes

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/112/8/3330.full.html>

Articles on similar topics can be found in the following Blood collections
[Neoplasia](#) (4217 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Copyright 2011 by The American Society of Hematology; all rights reserved.



Desirable performance characteristics for *BCR-ABL* measurement on an international reporting scale to allow consistent interpretation of individual patient response and comparison of response rates between clinical trials

Susan Branford,¹ Linda Fletcher,¹ Nicholas C. P. Cross,² Martin C. Müller,³ Andreas Hochhaus,³ Dong-Wook Kim,⁴ Jerald P. Radich,⁵ Giuseppe Saglio,⁶ Fabrizio Pane,⁷ Suzanne Kamel-Reid,⁸ Y. Lynn Wang,⁹ Richard D. Press,¹⁰ Kevin Lynch,¹¹ Zbigniew Rudzki,¹ John M. Goldman,¹² and Timothy Hughes¹

¹Institute of Medical and Veterinary Science, Adelaide, Australia; ²National Genetics Reference Laboratory, University of Southampton, Salisbury, United Kingdom; ³Medizinische Fakultät Mannheim, University of Heidelberg, Mannheim, Germany; ⁴St Mary's Hospital, The Catholic University of Korea, Seoul, Korea; ⁵Fred Hutchinson Cancer Research Center, Seattle, WA; ⁶Ospedale Università di Torino, Turin, Italy; ⁷Hematology Unit, Ceinge and Dipartimento di Biochimica e Biotecnologie Mediche, University of Naples Federico II, Naples, Italy; ⁸Princess Margaret Hospital, Toronto, ON; ⁹Weill Medical College of Cornell University, New York, NY; ¹⁰Oregon Health & Science University, Portland; ¹¹Novartis Pharmaceuticals Australia, Sydney, Australia; and ¹²Imperial College at Hammersmith Hospital, London, United Kingdom

An international basis for comparison of *BCR-ABL* mRNA levels is required for the common interpretation of data derived from individual laboratories. This will aid clinical decisions for individual patients with chronic myeloid leukemia (CML) and assist interpretation of results from clinical studies. We aligned *BCR-ABL* values generated by 38 laboratories to an international scale (IS) where a major molecular response (MMR) is 0.1% or less. Alignment was achieved by application of labo-

ratory-specific conversion factors calculated by comparisons performed with patient samples against a reference method. A validation procedure was completed for 19 methods. We determined performance characteristics (bias and precision) for consistent interpretation of MMR after IS conversion. When methods achieved an average *BCR-ABL* difference of plus or minus 1.2-fold from the reference method and 95% limits of agreement within plus or minus 5-fold, the MMR

concordance was 91%. These criteria were met by 58% of methods. When not met, the MMR concordance was 74% or less. However, irrespective of precision, when the bias was plus or minus 1.2-fold as achieved by 89% of methods, there was good agreement between the overall MMR rates. This indicates that the IS can deliver accurate comparison of molecular response rates between clinical trials when measured by different laboratories. (Blood. 2008;112:3330-3338)

Introduction

Serial analysis of *BCR-ABL* mRNA levels by real-time quantitative PCR (RQ-PCR) accurately reflects the level of leukemic inhibition induced by therapy and provides an appropriate monitoring strategy for patients with chronic myeloid leukemia (CML).¹⁻⁷ Many consider the achievement of a major molecular response (MMR) a goal of imatinib therapy because it is associated with a favorable progression-free survival.⁸⁻¹³ A recent report has recognized that failure to achieve an MMR by 18 months of imatinib treatment may represent a suboptimal response and review of therapy was recommended.¹⁴ Therefore, accurate molecular analysis is useful for a clinician to make informed patient management decisions.

The trial that established MMR as a clinically relevant end point was the IRIS study.^{8,9} For this trial, 3 testing laboratories standardized their RQ-PCR methods using a standardized baseline, and MMR was a 3-log reduction from this level.⁸ Establishment of a value for MMR in other laboratories equivalent to that established in the IRIS study has not been straightforward. Clearly, a process for accurate alignment of *BCR-ABL* values in each laboratory to the original MMR value will aid in the consistent measurement of MMR throughout the world. Serial molecular analysis within a laboratory will certainly provide benefit for individual patients, but alignment of *BCR-ABL* values is necessary for several reasons: the use of common clinical decision values, facilitation of patient

mobility between clinics that use different testing laboratories, and consistent interpretation of clinical research data, including those to be considered by regulatory authorities. One process to align data has been undertaken between the IRIS participating laboratory of Hughes and Branford in Australia and the laboratory of Hochhaus and Müller in Germany. A *BCR-ABL/ABL%* value in the German laboratory equivalent to the MMR value in the Australian laboratory was calculated by an exchange of patient samples. An association was demonstrated between the log reduction scale for determining MMR and a quantitative value using different methods and control genes.¹⁵

The existing RQ-PCR methods in use around the world use different techniques and various control genes, leading to marked variation in reported *BCR-ABL* values.¹⁶ The best approach for achieving consistent and comparable quantitative data on a global scale would be to use internationally established reference reagents.¹⁷ Work is in progress toward developing such material for *BCR-ABL* quantitation,¹⁸ but until they are available, an alternative approach to achieve comparable *BCR-ABL* values is the use of an international reporting scale.¹⁹ This scale is anchored to the value that defines an MMR as established by the IRIS study laboratories.⁸ The proposed international scale (IS) is designed to replace the log reduction scale with a defined value for MMR of 0.10% IS.¹⁹

Submitted April 9, 2008; accepted July 21, 2008. Prepublished online as *Blood* First Edition paper, August 6, 2008; DOI 10.1182/blood-2008-04-150680.

The publication costs of this article were defrayed in part by page charge

payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2008 by The American Society of Hematology

Conversion to the IS is achieved by the application of laboratory-specific conversion factors. Specific conversion factors are required for each laboratory because differences in the components that comprise the complete RQ-PCR analytical system contribute to variation in the reported values.²⁰

We present here results of an international collaborative study to test the feasibility of reporting *BCR-ABL* values on the IS by the application of laboratory-specific conversion factors that were derived using patient samples. The validity of the conversion factors was checked with analysis of subsequent samples. Alignment of data was demonstrated for most laboratories and performance characteristics that produce desirable concordance between laboratories at the critical decision level of MMR were identified. The process allows for differences generated by various RQ-PCR analytical systems. We anticipate that the project will lead to and guide the preparation of certified international standards that will be available to all laboratories.

Methods

Reference laboratory

The reference laboratory in Adelaide, Australia performed molecular analyses for patients enrolled in the IRIS study, and the quantitative value representing MMR (3-log reduction from the standardized baseline) is *BCR-ABL/BCR%* 0.08%.⁸ The RQ-PCR method has been detailed previously.^{2,21} To convert *BCR-ABL* values to IS in this reference laboratory, all values were multiplied by the conversion factor of 1.25. This conversion factor was calculated by dividing the value representing MMR IS (0.10%) by the quantitative value representing a 3-log reduction in this laboratory (0.08%). For the current study, the Adelaide reference laboratory coordinated the exchange of samples, analyzed all samples, and calculated and validated the conversion factors for the participating laboratories.

Participating laboratories

The RQ-PCR analytical systems tested for their suitability for IS conversion were those routinely used in 38 laboratories from 15 countries.^{1-3,8,11,22-25} Two laboratories used 2 different methods that varied in the control gene; therefore, 40 analytical systems were tested. The analytical system of each laboratory comprised the control gene, the primer and probe sequence location, the probe chemistry, the material used for the construction of standard curves, the instrumentation, the reverse transcription procedure, the quantitative PCR procedure, all of the reaction components, and the operator technique. These components contribute to variability between the methods. Table 1 summarizes the components of each analytical system and indicates that there was considerable variation in the techniques. For purposes of this study, the analytical systems of the participating laboratories are called field methods. Laboratory-specific conversion factors for each field method are designed to overcome the potential differences in reported *BCR-ABL* values due to different instruments, probe technologies, and other analytical processes.

Calculation of the field method-specific conversion factors

Each participating laboratory sent a median of 20 patient samples, either RNA (n = 431) or cells stored in Trizol reagent (Invitrogen, Carlsbad, CA) on dry ice (n = 353) to the reference laboratory for RQ-PCR analysis. The samples included those of patients in various disease phases and trials and were first analyzed in the originating laboratory as part of the usual molecular monitoring practice of the laboratory. Therefore, the samples were analyzed in the originating laboratory over several weeks or months and several RQ-PCR runs. The quantitative values generated by each field method for their set of patient samples was withheld from the reference laboratory until analysis in the reference laboratory was complete. Only samples with the common b2a2 (e13a2) and/or b3a2 (e14a2) *BCR-ABL*

Table 1. Details of the 40 field methods and the reference laboratory method

Characteristic	Number
Regions where the methods were performed	
Australia/New Zealand	13
Asia	10
North America	8
Europe	8
South America	1
South Africa	1
Reference for the RQ-PCR methods	
Gabert et al, ²² Beillard et al ²³ (EAC)	17
Branford et al ^{2,21}	12
Emig et al ¹	3
Emig et al, ¹ Müller et al ²⁵	1
Radich et al, ³ Hughes et al ⁸	1
Press et al ¹¹	1
Stock et al ²⁴	1
In house	5
Instruments	
Applied Biosystems	27
LightCycler	10
Corbett Rotorgene	4
Control genes	
<i>ABL</i>	18
<i>BCR</i>	15
<i>GUS</i>	4
<i>G6PDH</i>	3
β 2M	1
RNA extraction	
Trizol reagent	25
Qiagen column-based methods	10
MagnaPure LC mRNA isolation kit	2
PAXgene stabilization	1
RNA Stat-60	1
TRI reagent solution	1
Versagene RNA purification	1
Reverse transcriptase	
Superscript	24
MMLV	11
Roche Transcriptor	3
AMV	2
Omniscript/Sensiscript	1
RT primer	
Random hexamers	38
Gene specific	2
Random hexamers/Oligo dT	1
Material used for standard curve	
Plasmid	31
RNA	5
cDNA	3
No standard curve	2

transcripts were included. The effective measurement range for the international scale was deemed to be a *BCR-ABL* level of 10% IS or below. This was because most field methods used *ABL* as the control gene. Depending on the PCR primer design for the *ABL* control, both wild-type *ABL* and *BCR-ABL* are amplified when *BCR-ABL* expression is high, and the ratio of *BCR-ABL/ABL%* could therefore underestimate the leukemic load.^{16,22} Similarly, *BCR* as the control may lead to overestimation of *BCR-ABL/BCR%* when *BCR-ABL* expression is high because normal cells have 2 wild-type alleles and *BCR-ABL*-expressing cells have 1. However, inaccurate ratios when *BCR-ABL* expression is above 10% IS may have minimal impact on the interpretation of the result.²² In the majority of cases, these values indicate lack of a major cytogenetic response (MCR \leq 35% Philadelphia chromosome).²⁶ The tests performed using patient samples

were approved by the institutional review board or equivalent body of each institution. The research was conducted in accordance with the Declaration of Helsinki.

The reference laboratory analyzed each sample received from the participating laboratories by duplicate reverse transcription and quantitative PCR analysis, which is their usual practice. For each set of samples, only a few samples were analyzed on any one day. Therefore, the complete sample set was analyzed over several weeks in several RQ-PCR runs to mimic as closely as possible the analysis in the originating laboratory. This was to account for day-to-day variations of the test, reagent batches, and operator, all of which contribute to variability of the analytical system. This procedure is in accordance with Food and Drug Administration (FDA)-approved guidelines for method comparison using patient samples.²⁷

To determine the conversion factors, each set of data generated by a field method was compared with that generated by the reference laboratory for the same sample set. Using the method comparison procedure of Bland and Altman, the bias between the field method and the reference method was determined.^{28,29} Briefly, this method plots the difference between the 2 measurements against their mean to indicate whether there is a systematic difference between the 2 measurements. For example, there may be a consistent tendency for a field method to exceed the reference method, which is the bias. The 95% limits of agreement provide an interval within which most of the individual differences can be expected to lie. This procedure has been used previously to measure the agreement between peripheral blood and bone marrow *BCR-ABL* values.²⁴ The estimate is applicable only if the bias is consistent across the measurement range. If this is the case, then alignment between the 2 methods can be achieved by applying a conversion factor.²⁹

For each field method, the first step in the conversion factor calculation was conversion of the *BCR-ABL* values generated by the reference laboratory to the IS by multiplying by 1.25 (Adelaide laboratory-specific conversion factor). These data and the original *BCR-ABL* values generated by the field method were \log_{10} -transformed. A bias plot was generated by plotting the difference between the 2 methods against the mean of the methods, with the field method data as the *X* variable. The 95% limits of agreement were estimated by mean difference plus or minus 1.96 SDs of the differences and provided an interval within which 95% of differences between measurements were expected to lie. The antilog of the estimated mean bias between the methods was designated as the conversion factor for the field method.

Validation of the laboratory-specific conversion factors

Once the conversion factors for each field method were calculated, they required validation. For this purpose, the participating laboratories sent a subsequent set of samples to the reference laboratory for analysis as described above. To date, 20 of the field methods have sent samples to the reference laboratory for the validation process (Australia/New Zealand *n* = 7, Europe *n* = 5, North America *n* = 4, and Asia *n* = 4). Validation samples from the remaining field laboratories are pending. The validation process involved an exchange of 458 RNA samples and 229 samples stored in Trizol reagent. Of these, 598 samples were suitable for validation of the conversion factors. The remaining samples were inappropriate because the RNA was degraded or the *BCR-ABL* values were above the IS effective measurement range. The validation samples were sent to the reference laboratory at a median of 7 months (range, 2-18 months) after the conversion factor for a field method was calculated. For 5 field methods, more than 1 set of validation samples were sent over several months. The reference laboratory analyzed each validation sample in the same manner as those sent for the conversion factor calculation. The *BCR-ABL* values of the reference laboratory and the corresponding field method for each set of validation samples were converted to the IS by multiplying by their specific conversion factors. The agreement between the field method and the reference method after conversion was assessed using the method of Bland and Altman.^{28,29}

Results

Relationship between the *BCR-ABL* international scale and cytogenetic response levels

BCR-ABL values converted to the IS equate approximately to cytogenetic response levels. In the Adelaide reference laboratory a comparison of peripheral blood *BCR-ABL* RQ-PCR values with their corresponding bone marrow cytogenetic assessment was undertaken for 828 samples.²⁶ The *BCR-ABL* values were converted to IS. In 98% of the *BCR-ABL* values of more than 1% to 10% IS (*n* = 142), the corresponding Philadelphia chromosome percentage was within the range of an MCR. Of the *BCR-ABL* values of 1.0% or less IS (*n* = 530), 96% correspondingly indicated a complete cytogenetic response (CCR) and 100% indicated an MCR. The 4% of samples with *BCR-ABL* values of 1.0% IS and below that correspondingly did not indicate CCR had a median Philadelphia chromosome percentage of 3%.

Validity of the reference laboratory to align data to international units

The IS is aligned to the value representing MMR as established in the IRIS study in 2001.⁸ For the Adelaide laboratory to be an appropriate reference laboratory, it must demonstrate that it can trace its MMR value directly to that originally established in the IRIS study in 2001. The specific *BCR-ABL/BCR%* value representing MMR in Adelaide is 0.08%. Therefore, the laboratory was required to demonstrate that a sample with a measured *BCR-ABL* value of 0.08% in 2001 would still be 0.08% (or within the measurement reliability of the analytical system) when measured today.

To ensure the consistent performance of the Adelaide RQ-PCR analytical system, quality control (QC) samples were included in every run.^{20,21,30} Briefly, 2 different QC samples with a high and low *BCR-ABL* level were processed in the same way as the patient samples in every run. These samples had pre-established *BCR-ABL* values and, for the run to be accepted, the values must be within a defined range that is based on the SD of the analytical system. Runs with QC sample values outside the range are rejected and the analysis repeated. The QC samples were prepared in large batches and frozen in aliquots of RNA stabilization solution. The target mean QC values were established before their introduction into routine use. The yearly mean quality control values since 2001 are detailed in Table 2. These demonstrate the high reproducibility of the Adelaide RQ-PCR analytical system, which allows traceability and comparison to the original IRIS MMR value.

A second consideration in establishing the validity of the reference method for patient sample comparison studies is the inherent method variation. A reference method should have not only minimal bias from the reference value, but also minimal imprecision.³¹ To determine the bias and limits of agreement of the reference method, patient samples were analyzed in duplicate within the reference laboratory. Duplicate analyses in this case involved splitting 163 blood samples (20 mL each) into 2 tubes followed by separate RNA extraction, reverse transcription, and quantitative PCR analysis for the duplicate blood samples on different days over approximately 18 months. The variables were the operator, day of analysis, reagent batches, and calibration status of the instrument and pipettes. The instrument, reverse transcription reaction, control gene, and RQ-PCR method did not vary. A bias plot between the first and second measurement was generated which showed negligible mean bias (Figure 1). The spread of

Table 2. Yearly mean and coefficient of variation (CV) values for the quality control (QC) samples of the Adelaide reference laboratory

Control	QC target values	2001	2002	2003	2004	2005	2006 (to March 28)	New QC target values	2006 (from March 29)	2007
Low b2a2										
Mean	0.07	0.07	0.07	0.09	0.08	0.06	0.07	0.06	0.06	0.06
CV, %		32	55	34	35	40	34		30	29
High b2a2										
Mean	56	48	53	69	50	42	52	22	21	20
CV, %		33	33	22	25	28	29		26	29
Low b3a2										
Mean	0.07	0.07	0.06	0.08	0.06	0.08	0.09	0.08	0.08	0.08
CV, %		54	52	37	41	35	34		34	35
High b3a2										
Mean	85	82	77	93	69	94	93	27	26	24
CV, %		26	38	20	18	23	31		26	22
Number of RQ-PCR runs		34	110	147	157	212	77		257	192

All values were recorded whether the run was accepted or rejected. A new batch of QC samples was in use from March 2006.

results as estimated by the 95% limits of agreement was plus or minus 2.5-fold of the mean, which is an indication of within-method variability. Overall, 90% of values generated by the second measurement were within plus or minus 2-fold of the first measurement. The within-method variability of the reference method indicated a baseline against which to judge between-method variability of the reference method and the field methods after IS conversion.

Conversion factors for each field method

Conversion factors were calculated for 36 of the 40 field methods. Samples from 4 field methods were inappropriate to calculate the conversion factor because the RNA was degraded or the *BCR-ABL* values were above the IS effective measurement range. The calculated conversion factors for the 36 field methods ranged from 0.18 to 13.5. Conversion factors differed even though several field methods used the same primer, probe and control gene combinations. For example, for 14 field methods using the Europe Against Cancer (EAC) primer and probe set and the *ABL* control gene,^{22,23} the conversion factors ranged from 0.23 to 1.40. This indicated that the estimated mean bias of these methods ranged from 1.4-fold lower to 4.3-fold higher than the converted reference method data. For 11 field methods using a common primer and probe set and the *BCR* control gene,² the conversion factors ranged from 0.42 to

5.27. This indicated that the estimated mean bias of these methods ranged from 5.3-fold lower to 2.4-fold higher than the converted reference method data. Differences in the conversion factors were also evident among the Australian laboratories despite having the opportunity to consult more closely with the reference laboratory. The variability of the conversion factors emphasizes that it is the complete analytical system that contributes to variation of data. Figure 2A and C demonstrate the conversion factor calculation process of one field method. This is representative of the plots

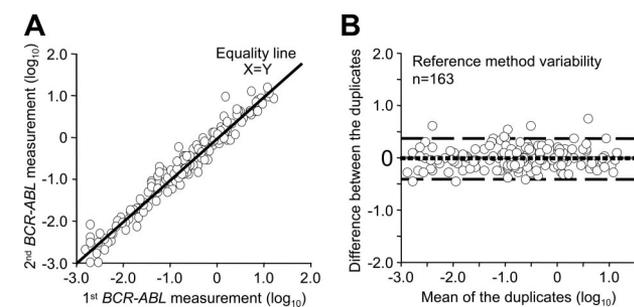


Figure 1. Reference method variability. A bias plot was generated to determine the variability of the reference method when patient samples were measured in duplicate. (A) Visual inspection of the data showed that both measurements closely follow the equality line. (B) On the x-axis of the bias plot is the mean value of the duplicates and on the y-axis is the difference between the 2 measurements. The dotted line is the estimated mean bias from 0 between the measurements. If there was no difference in the mean bias this line would be at 0, which it is. The dashed lines are the 95% limits of agreement, which indicates that 95% of the second measurements were within approximately plus or minus 2.5-fold of the first measurement. Ninety percent of the values were within 2-fold.

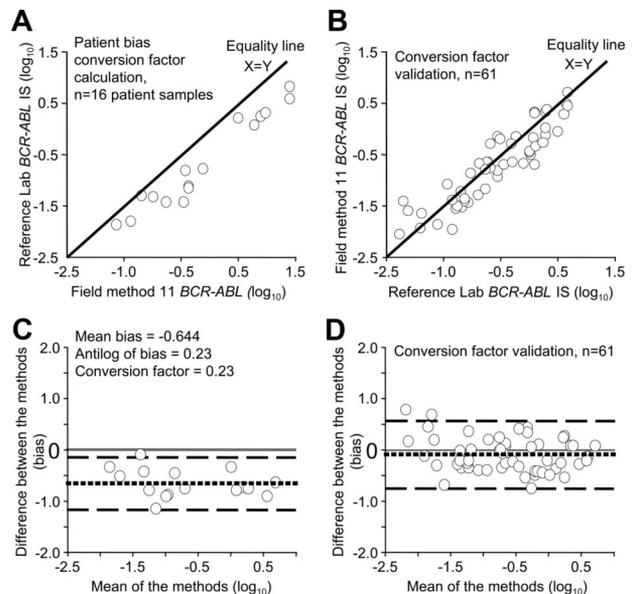


Figure 2. Conversion factor calculation and validation for field method 11. (A) Visual inspection of the data used for the conversion factor calculation showed that all values generated by field method 11 were higher than the reference method, indicating a consistent bias between the methods. (C) Bias plot of the same data. If there were no difference in the estimated mean bias, the dotted line would be at 0. The values generated by field method 11 were, on average, 4.3-fold higher than the values generated by the reference method. The conversion factor of 0.23 was the antilog of the mean bias. The conversion factor was validated by subsequent sample exchange. Visual inspection of the converted validation data (B) and a plot of the bias (D) indicated the estimated mean bias was close to 0. The field method generated values that were on average 1.2-fold lower than the reference method, indicating a considerable improvement in data alignment. It was estimated that 95% of values were within the range of 4.9-fold lower to 4.5-fold higher than the reference method. Significantly, 59% of all samples were within plus or minus 2-fold of the reference value. Before conversion only 14% were within 2-fold. The conversion factor calculation required the X variable to be the field method. However, for the validation process the reference laboratory data were the X variable to determine the relative difference.

generated for all field methods, however there was one field method that did demonstrate markedly inconsistent bias (values up to 10-fold lower at the lower measurement limit and 10-fold higher at the upper measurement limit). However, insufficient samples were received from this laboratory to allow a true assessment of bias, and further samples were requested.

Validation of the conversion factors

The reliability of the conversion factor for each field method to align data to the IS was dependent on the consistency of analysis within each laboratory. This means that the same value should be reproduced for a particular sample, within the calculated variability of that assay, when measured over various time points. Twenty field methods sent samples to validate their conversion factors, although only 19 have completed the process for their current method. The bias between each field method and the reference method was calculated before and after conversion to the IS using the specific conversion factor of each method. Figure 2B,D demonstrates the bias plot of one field method after conversion to the IS.

The estimated mean bias of each field method after conversion was calculated as the average fold difference compared with the reference method. An average difference of 1.0-fold would indicate that there was no difference in the average *BCR-ABL* IS values. For 16 of the 20 field methods, the average difference was plus or minus 1.2-fold. Before conversion, only 5 field methods had an average difference of plus or minus 1.2-fold. The remainder ranged from 7.7-fold lower to 8.1-fold higher than the reference method. After conversion, the estimated 95% limits of agreement for each method varied. The method with the closest agreement to the reference method had a 95% range of plus or minus 2.7-fold. The method with the least agreement had a 95% range of approximately plus or minus 8-fold (Table 3).

Three of the field methods still showed a consistent bias after conversion where almost all values were either greater than or less than the reference method values. This was indicated by an average difference of +1.8-fold, +2.1-fold and -2.7-fold for the 3 methods.

Two of the methods had altered one of the components of their analytical system from the time of the conversion factor calculation to the time of validation. One had optimized the random primer concentration in accord with the recommendations of the EAC,²³ and the other had changed the reverse transcriptase enzyme. The conversion factors for these 2 field methods required recalculation (one changed from 2.02 to 1.13 and the other from 1.65 to 4.1). This demonstrates that seemingly minor alterations to an analytical system may have a significant impact on the measurement. One of the methods has been revalidated using the new conversion factor and the average difference was -1.2-fold. Therefore, 17 of 19 field methods achieved an average difference of plus or minus 1.2-fold for their current method after the validation process.

Figure 3 demonstrates the average difference and the 95% limits of agreement before and after IS conversion for 19 field methods. These 19 methods have undertaken the validation procedure for their current method. One laboratory used both *β2M* and *ABL* as control genes in 2 separate assays. Therefore, the methods within this laboratory were considered as 2 separate field methods (Table 3, field methods 3 and 7). Before conversion the average differences for these methods were -7.2-fold for the *β2M* method and +8.1-fold for the *ABL* method. This demonstrates that a single major variable such as the control gene can have a significant impact on the measurement. After application of the field method-specific conversion factors for this laboratory, the average difference was +1.1 and +1.2-fold, respectively, compared with the reference method.

Evaluation of the agreement between methods after IS conversion

When judging acceptable agreement between the methods after IS conversion one must consider what constitutes acceptable accuracy (average fold difference relative to the reference method average value) and precision (reproducibility as assessed by the 95% limits of agreement) for appropriate clinical decisions. These could be considered as the performance characteristics that optimally assign

Table 3. Summary of performance characteristics in terms of the differences in *BCR-ABL* values compared to the reference method of the 19 field methods that have undertaken the conversion factor validation procedure for their current analytical system

Field method	Control gene	RQ-PCR reference	Conversion factor	Number of validation samples	Average difference before conversion (fold)	Average difference after conversion (fold)	95% limits of agreement after conversion	
							Lower (fold)	Upper (fold)
1	<i>ABL</i>	22,23	1.35	21 trizol	-1.1	1.0	-2.7	+2.7
2	<i>BCR</i>	2	1.13	26 trizol	-1.1	-1.2	-2.7	+2.3
3	<i>β2M</i>	In house	10.23	17 RNA	-7.2	+1.1	-2.4	+2.6
4	<i>BCR</i>	2	1.7	16 trizol	+1.1	+1.2	-2.5	+2.9
5	<i>BCR</i>	2	1.05	17 trizol	+1.0	-1.2	-3.3	+2.9
6	<i>BCR</i>	2	1.28	58 RNA	-1.1	-1.1	-3.5	+3.3
7	<i>ABL</i>	In house	0.18	14 RNA	+8.1	+1.2	-3.4	+3.8
8	<i>ABL</i>	22,23	0.56	20 trizol	+1.6	+1.2	-3.5	+3.9
9	<i>ABL</i>	1	0.88	26 RNA, 22 trizol	+1.4	1.0	-4.3	+4.3
10	<i>BCR</i>	In house	10.2	20 RNA	-6.0	-1.2	-4.6	+4.2
11	<i>ABL</i>	22,23	0.23	61 RNA	+4.4	-1.2	-4.9	+4.5
12	<i>ABL</i>	In house	7.78	15 RNA	-7.7	-1.2	-5.2	+4.8
13	<i>ABL</i>	22,23	0.36	22 trizol	+4.4	+1.2	-5.2	+4.8
14	<i>ABL</i>	22,23	0.56	73 RNA	+2.9	+1.3	-4.5	+5.1
15	<i>ABL</i>	22,23	0.79	32 RNA	+1.7	+1.2	-5.8	+5.9
16	<i>BCR</i>	3,8	2.39	61 RNA, 32 trizol	-2.0	-1.2	-6.1	+5.9
17	<i>BCR</i>	In house	0.42	14 trizol	+2.5	-1.2	-7.0	+6.6
18	<i>GUS</i>	1,25	2.14	17 trizol	-1.7	+1.2	-7.6	+8.0
19	<i>GUS</i>	22,23	0.61	14 trizol	+4.2	+2.1	-2.7	+4.9

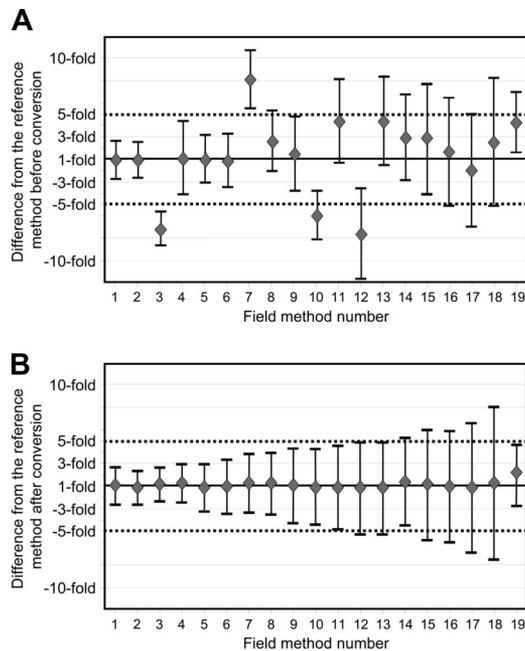


Figure 3. Average difference and 95% limits of agreement between the reference method and the field methods before and after conversion to the IS. Nineteen field methods have been submitted to the validation procedure for their current analytical system. (A) Before conversion, the average difference between each field method and the reference method as represented by the diamonds ranged from 7.7-fold lower to 8.1-fold higher. A difference of 1.0 indicates no difference between the average of the *BCR-ABL* values. The bars above and below the average difference define the 95% limits of agreement for that field method. (B) After conversion, the average differences were within plus or minus 1.2-fold for most field methods. However, the 95% limits of agreement varied.

MMR for samples, while acknowledging that the inherent within-method variability of the current technology may be at least plus or minus 2-fold in approximately 90% of samples. Performance characteristics were defined as the average fold difference and the 95% limits of agreement after conversion. Seventeen of the 19 field methods (89%) were able to achieve an average difference of plus or minus 1.2-fold after conversion. Of these 17 field methods, for 11 (group 1) the 95% limits of agreement were also within plus or minus 5-fold, whereas for 6 field methods (group 2) the 95% limits of agreement were greater than plus or minus 5-fold (Table 3). The remaining 2 field methods were designated as group 3. The concordance of MMR IS was determined between the reference method and the 3 groups. Concordance of MMR means the reference method generated a *BCR-ABL* value of 0.10% or less of IS for a particular sample and the corresponding value for that sample generated by the field method was also 0.10% or less IS. Table 4 demonstrates that group 1 achieved the best concordance at 91%. In this group, 14 of the 15 discordant values (93%) were within 2-fold of the upper MMR range. For our analytical procedure, the inherent within-method variability for 90% of

samples is 2-fold; therefore, the discordance of MMR for these 11 group 1 field methods is indistinguishable from the inherent assay variability. Therefore, the MMR concordance rate of 91% may be close to optimal with current technology. The rate of MMR concordance for group 1 was much better than that seen for group 2 (74%) and group 3 (60%).

As indicated in Table 3, the 11 field methods that produced the best concordance of MMR were diverse. They were performed in 7 countries and used various control genes, instrument platforms, and reagents. These 11 methods had less variability for this particular study using the samples available, and it should not be inferred that the other methods may not achieve these performance characteristics under optimized study conditions.

Discussion

We have demonstrated that variation between *BCR-ABL* values generated by diverse analytical systems can be substantially reduced by aligning data on an international reporting scale using method-specific conversion factors. Bias plots were used to calculate and validate conversion factors, since these have been recommended as an acceptable procedure for method comparison studies using clinical samples.^{27,32} Although the method of the Adelaide laboratory was used as the reference method for this study, it should not be inferred that data generated by this method (or any reference method) are without error in terms of variability.²⁹ Some lack of agreement between different methods of measurement is inevitable. What matters is the amount by which methods disagree and whether the differences will cause problems in clinical interpretation.²⁹ We identified performance characteristics that limited the discordance of MMR IS, a molecular value which is used for clinical decisions as based on the recommendations of a panel of experts¹⁴ and is also the primary end point of clinical trials of ABL inhibitors.^{33,34} These performance characteristics were an average difference of plus or minus 1.2-fold between the reference method and a field method, which indicated the bias between methods, and 95% limits of agreement of plus or minus 5-fold of the reference method, which indicated the precision or a measurement of the inter-assay variability. These performance characteristics were achieved by 58% of the participants, leading to an MMR concordance of 91%. This indicates that less than 10% of patients would be misclassified when these performance criteria are met. The samples that were misclassified had *BCR-ABL* values within 2-fold of the upper limit of MMR in 93% of cases, which is consistent with the inherent within-method variability. The MMR concordance rate of 91% may, therefore, be the maximum that can be achieved using the current technology.

It is important to define analytical goals for bias and precision to limit diagnostic misclassifications.^{35,36} This is particularly important for clinical situations in which a diagnostic test is used for

Table 4. Concordance of MMR IS between reference method and field methods according to the performance characteristics of the 19 field methods that have undertaken the validation procedure for their current analytical systems

Field methods group	Performance characteristics		Number of field methods (%)	Number of validation samples	MMR by either reference method and/or field methods	MMR by reference method (% of all validation samples)	MMR by field methods (% of all validation samples)	MMR concordance (%)
	Average difference, ±	95% limits of agreement						
1	1.2-fold	±5-fold	11 (58)	318	159	152 (48)	151 (47)	144 (91)
2	1.2-fold	> ±5-fold	6 (32)	193	94	83 (43)	79 (41)	70 (74)
3	> 1.2-fold	any	2 (10.5)	87	25	22 (25)	17 (20)	15 (60)

single-point classification of patients, as it is for molecular testing in CML. The primary end point of various current studies of imatinib and second-generation ABL kinase inhibitors is the rate of MMR, which is frequently determined in different molecular laboratories using diverse analytical procedures. To enable comparison of response rates across studies, it is essential that the molecular data are reliable. Increasing the analytical bias leads to a higher level of diagnostic misclassification,³⁷ and we demonstrated in this study that increasing the analytical imprecision while maintaining bias increased the incidence of misclassification of MMR for individual samples from 9% to 26%. However, when the bias was maintained within plus or minus 1.2-fold of the reference method, which was the case for 89% of the field methods, the imprecision did not unduly influence the overall rates of MMR between the reference method and the field methods. When the 95% limits of agreement were within plus or minus 5-fold of the reference method, 48% of all samples were MMR as measured by the reference method versus 47% by the field methods (Table 4). When the bias was plus or minus 1.2-fold but the 95% limits of agreement were greater than plus or minus 5-fold, 43% of all samples were MMR as measured by the reference method versus 41% by the field methods. This indicates that when the bias is minimized there is good agreement between the overall rates of MMR when different laboratories undertake the molecular analysis. Therefore, reporting *BCR-ABL* values on the IS will allow accurate comparison of response rates for clinical trials, providing, however, that laboratories maintain the consistent performance of their RQ-PCR analytical system. This can be monitored by the inclusion of QC samples in every run to recognize shifts in data, which should be corrected before results are deemed acceptable.^{20,21,30} Otherwise, the laboratory-specific conversion factor may be invalid.

The performance characteristics of all field methods for this study may not necessarily reflect their actual performance characteristics, since the number of samples available for analysis varied. This may influence the conversion factor calculation and validation, and hence the assessment of analytical performance.³⁸ The reliability and effectiveness of the comparison increases with analysis of more samples over more time.²⁷ Furthermore, assessment of precision should ideally be performed using stable QC samples with defined *BCR-ABL* values that are analyzed at least 10 times over several runs. Such material has currently not been tested and is not available. Therefore, the procedures used in this study require refinement to appropriately assess the performance of each laboratory. Nevertheless, we have established that limiting the variability within an analytical system is important for reliable clinical interpretation. Furthermore, the appropriate interpretation of a rise in *BCR-ABL* level is dependent on knowledge of the variability of an RQ-PCR field method.^{13,20,39,40} For an assay with greater variability, a change of more than 5- to 10-fold in *BCR-ABL* level may be clinically significant, whereas for an assay with less variability a 2- to 3-fold change may be clinically significant. Precision can be improved by adherence to accepted principles of good laboratory practice and quality assurance.

We anticipate that international reference materials, when widely available, will eventually replace the method used in this study to align data. Candidate materials are under development and production of an initial batch is scheduled for late 2008. In addition, it is likely that one or more commercial companies will also be producing calibration reagents that will need to be aligned to the IS. However, validation of these materials will take several months and, even assuming they perform well (which is not

guaranteed), the accreditation process will take longer. Therefore, there will be a need to continue the conversion factor process for some time to achieve standardization. This may be a formidable task for laboratories where *BCR-ABL* analysis is a minor component of their testing procedures.

The preparation of reference material may be a complex issue, as the matrix must closely mimic that of patient samples. Calibrating a matrix-sensitive analytical system with reference reagents that react differently from patient samples leads to significant biases and inaccurate values when patient samples are analyzed.⁴¹ The international reference materials will also need to account for all of the variability of the current RQ-PCR analytical systems. Müller et al²⁵ demonstrated, in a study involving 37 laboratories, that even with a common plasmid standard and optimized methods there was still considerable variation in reported *BCR-ABL* values. When *ABL* was used as the control gene, *BCR-ABL* values generated by methods using TaqMan platforms were, on average, 1.7-fold higher than laboratories using LightCycler platforms. Similarly, when *GUS* was used as the control gene the values were 2.2-fold higher for TaqMan platforms than laboratories using LightCycler platforms. Furthermore, we have demonstrated in this study that a change of a single step in the reaction process can substantially shift *BCR-ABL* values. Therefore, laboratories may require realignment to the IS when any aspect of the analytical procedure is changed. If a single system for measuring *BCR-ABL* is to be introduced in the future, it must overcome the inherent variability within each laboratory that contributes to differences in the reported value. It remains to be determined whether *BCR-ABL* values generated by cartridge-based microfluidic systems that incorporate RNA extraction, reverse transcription, and quantitative PCR^{42,43} will be interchangeable between laboratories.

This study was not designed to allow an assessment of any difference in sensitivity between the methods. Furthermore, we could not appropriately assess variability of the reported value that may be associated with the RNA extraction process in some cases. To address these issues, the samples used should have guaranteed stability and allow individual laboratories to extract RNA using their usual procedure. Matrix effects must be considered and excluded for any material used for this purpose.

Undoubtedly, more robust statistical approaches will be required in the long term to determine conversion factors for the IS. More than one conversion factor may be required to adequately report values over the wide measurement range. Nevertheless, this study has provided a proof of principle that alignment of *BCR-ABL* values generated from diverse methods is achievable using an international reporting scale. We have identified desirable performance characteristics to allow consistent measures of drug response that can be traced to published outcomes and recommendations and that facilitate the interpretation of national and international clinical studies.

Acknowledgments

The authors acknowledge the contribution of S. Armitage and D. Fairbairn, Royal Brisbane Hospital, Brisbane, Australia; A. Bell, Royal Melbourne Hospital, Melbourne, Australia; I. Bendit, Prédio dos Ambulatórios, Sao Paulo, Brazil; L. Beppu-Wong, Fred Hutchinson Cancer Research Center, Seattle, WA; A. Chan and Yip Sze Fai, Queen Mary Hospital, Hong Kong; K. Cheng, Prince of Wales Hospital, Chinese University of Hong Kong; C. Chuah and G.F. How, Singapore General Hospital, Singapore; D. Colomer,

Hospital Clinic of Barcelona, Spain; S. Ebrahim, Inkosi Albert Luthuli Central Hospital, Durban, South Africa; H. Goh, St Mary's Hospital, The Catholic University of Korea, Seoul, Korea; V. Hanrahan, Canterbury Health Laboratories, Christchurch, New Zealand; M. Hertzberg and D. MacDonald, Westmead Hospital, Sydney, Australia; M. Higgins, Royal Perth Hospital, Perth, Australia; H. Iland and A. Catalano, Kanematsu Laboratories, University of Sydney, Sydney, Australia; T. Zhang, Princess Margaret Hospital, Toronto, ON; E. Koay, National University Hospital, Singapore; L. Kann, Genzyme Genetics; S. Langabeer, St James Hospital, Dublin, Republic of Ireland; J.H. Liu, Taipei Veterans General Hospital, Taipei, Taiwan; E. Ma, Hong Kong Sanatorium & Hospital, Hong Kong; G. Martinelli, Institute of Hematology and Medical Oncology, Bologna, Italy; M. McBean and S. Kovalenko, Peter MacCallum Cancer Institute, Melbourne, Australia; N. Pattle, Doravitch Gippsland Pathology, Victoria, Australia; F. Quarantelli, University of Naples Federico II, Naples, Italy; L. Rozen and H. El Housni, Hospital Erasme, Brussels, Belgium; L.Y. Shih, Chang Gung Memorial Hospital, Taoyuan Hsien, Taiwan; W. Stevenson, Royal North Shore Hospital, Sydney, Australia; W. Stock and D. Sher, University of Chicago, Chicago, IL; J.L. Tang, National Taiwan University Hospital, Taipei, Taiwan; D. Taylor, Mater Hospital, Brisbane, Australia; N. Van De Water, Auckland Hospital, Auckland, New Zealand; and M. Wong, Tuen Mun Hospital, Hong Kong.

We thank Rebecca Lawrence and the staff of the Molecular Pathology, Institute of Medical and Veterinary Science, Adelaide, Australia, for assistance in analyzing the samples and all of the other laboratories who have participated in this study subsequent to the current data analysis.

This study was supported in part by research funding from Novartis Pharmaceuticals.

Authorship

Contribution: S.B. designed the research, interpreted the data and wrote the manuscript; L.F. performed and analyzed experiments; T.H. designed the research, interpreted the data and contributed to writing the manuscript; and Z.R., N.C., M.M., A.H., J.R., D.K., G.S., F.P., S.K.R., Y.L.W., R.P., K.L., and J.G. contributed to the interpretation of the data and writing the paper.

Conflict-of-interest disclosure: K.L. was an employee of Novartis Pharmaceuticals. S.B., T.H., N.C., M.M., A.H., J.R., D.K., G.S., F.B., and J.G. received research grants and/or honoraria from Novartis Pharmaceuticals. The remaining authors declare no competing financial interests.

Correspondence: Susan Branford, Division of Molecular Pathology, Institute of Medical and Veterinary Science, Adelaide, South Australia, 5000, Australia; e-mail: susan.branford@imvs.sa.gov.au.

References

- Emig M, Saussele S, Wittor H, et al. Accurate and rapid analysis of residual disease in patients with CML using specific fluorescent hybridization probes for real time quantitative RT-PCR. *Leukemia*. 1999;13:1825-1832.
- Branford S, Hughes TP, Rudzki Z. Monitoring chronic myeloid leukaemia therapy by real-time quantitative PCR in blood is a reliable alternative to bone marrow cytogenetics. *Br J Haematol*. 1999;107:587-599.
- Radich JP, Gooley T, Bryant E, et al. The significance of bcr-abl molecular detection in chronic myeloid leukemia patients "late," 18 months or more after transplantation. *Blood*. 2001;98:1701-1707.
- Wang L, Pearson K, Pillitteri L, Ferguson JE, Clark RE. Serial monitoring of BCR-ABL by peripheral blood real-time polymerase chain reaction predicts the marrow cytogenetic response to imatinib mesylate in chronic myeloid leukaemia. *Br J Haematol*. 2002;118:771-777.
- Kim YJ, Kim DW, Lee S, et al. Comprehensive comparison of FISH, RT-PCR, and RQ-PCR for monitoring the BCR-ABL gene after hematopoietic stem cell transplantation in CML. *Eur J Haematol*. 2002;68:272-280.
- Lee WI, Kantarjian H, Glassman A, Talpaz M, Lee MS. Quantitative measurement of BCR/abl transcripts using real-time polymerase chain reaction. *Ann Oncol*. 2002;13:781-788.
- Müller MC, Gattermann N, Lahaye T, et al. Dynamics of BCR-ABL mRNA expression in first-line therapy of chronic myelogenous leukemia patients with imatinib or interferon alpha/ara-C. *Leukemia*. 2003;17:2392-2400.
- Hughes TP, Kaeda J, Branford S, et al. Frequency of major molecular responses to imatinib or interferon alfa plus cytarabine in newly diagnosed chronic myeloid leukemia. *N Engl J Med*. 2003;349:1423-1432.
- Druker BJ, Guilhot F, O'Brien SG, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N Engl J Med*. 2006;355:2404-2417.
- Cortes J, Talpaz M, O'Brien S, et al. Molecular responses in patients with chronic myelogenous leukemia in chronic phase treated with imatinib mesylate. *Clin Cancer Res*. 2005;11:3425-3432.
- Press RD, Love Z, Tronnes AA, et al. BCR-ABL mRNA levels at and after the time of a complete cytogenetic response (CCR) predict the duration of CCR in imatinib mesylate-treated patients with CML. *Blood*. 2006;107:4250-4256.
- Iacobucci I, Saglio G, Rosti G, et al. Achieving a major molecular response at the time of a complete cytogenetic response (CCgR) predicts a better duration of CCgR in imatinib-treated chronic myeloid leukemia patients. *Clin Cancer Res*. 2006;12:3037-3042.
- Press RD, Galderisi C, Yang R, et al. A half-log increase in BCR-ABL RNA predicts a higher risk of relapse in patients with chronic myeloid leukemia with an imatinib-induced complete cytogenetic response. *Clin Cancer Res*. 2007;13:6136-6143.
- Baccarani M, Saglio G, Goldman J, et al. Evolving concepts in the management of chronic myeloid leukemia: recommendations from an expert panel on behalf of the European LeukemiaNet. *Blood*. 2006;108:1809-1820.
- Paschka P, Branford S, Lorentz C, Hehlmann R, Hughes T, Hochhaus A. Comparison of "log reduction from median pretherapeutic value" vs ratio Bcr-Abl/Abi to express the therapeutic response in CML [abstract]. *Blood*. 2004;104:287a.
- Zhang T, Grenier S, Nwachukwu B, et al. Interlaboratory comparison of chronic myeloid leukemia minimal residual disease monitoring: summary and recommendations. *J Mol Diagn*. 2007;9:421-430.
- Robertson JS. International standardization of gene amplification technology. *Biologicals*. 1998;26:111-113.
- Cross NCP, Hughes TP, Hochhaus A, Goldman JM. International standardisation of quantitative real-time RT-PCR for BCR-ABL. *Leukemia Res*. 2008;32:505-506.
- Hughes T, Deininger M, Hochhaus A, et al. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood*. 2006;108:28-37.
- Branford S, Cross NCP, Hochhaus A, et al. Rationale for the recommendations for harmonizing current methodology for detecting BCR-ABL transcripts in patients with chronic myeloid leukaemia. *Leukemia*. 2006;20:1925-1930.
- Branford S, Hughes T. Diagnosis and monitoring of chronic myeloid leukemia by qualitative and quantitative RT-PCR. In: Iland H, Hertzberg M, Marlton P, eds. *Myeloid Leukemia: Methods and Protocols*. Methods in Molecular Medicine, vol. 125. Totowa, NJ: Humana Press;2006:69-92.
- Gabert J, Beillard E, van der Velden VH, et al. Standardization and quality control studies of "real-time" quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia: a Europe Against Cancer program. *Leukemia*. 2003;17:2318-2357.
- Beillard E, Pallisaard N, van der Velden VH, et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using "real-time" quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR): a Europe against cancer program. *Leukemia*. 2003;17:2474-2486.
- Stock W, Yu D, Karrison T, et al. Quantitative real-time RT-PCR monitoring of BCR-ABL in chronic myelogenous leukemia shows lack of agreement in blood and bone marrow samples. *Int J Oncol*. 2006;28:1099-1103.
- Müller MC, Erben P, Saglio G, et al. Harmonization of BCR-ABL mRNA quantification using a uniform multifunctional control plasmid in 37 international laboratories. *Leukemia*. 2008;22:96-102.
- Ross DM, Branford S, Moore S, Hughes TP. Limited clinical value of regular bone marrow cytogenetic analysis in imatinib-treated chronic phase CML patients monitored by RQ-PCR for BCR-ABL. *Leukemia*. 2006;20:664-670.
- Krouwer JS, Tholen DW, Garber CC, et al. Method Comparison and Bias Estimation Using

- Patient Samples; Approved Guideline—second edition. (NCCLS document EP9-A2.) Wayne, PA: Clinical and Laboratory Standards Institute, 2002.
28. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
 29. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135-160.
 30. Hughes TP, Branford S. Monitoring disease response. In: Melo JV, Goldman JM, eds. *Myeloproliferative Disorders*. Berlin: Springer-Verlag; 2007:143-164.
 31. Bennett ST, Eckfeldt JH, Belcher JD, Connelly DP. Certification of cholesterol measurements by the National Reference Method Laboratory Network with routine clinical specimens: effects of network laboratory bias and imprecision. *Clin Chem*. 1992;38:651-657.
 32. Hollis S. Analysis of method comparison studies. *Ann Clin Biochem*. 1996;33(Pt 1):1-4.
 33. Cortes J, O'Brien S, Jabbour E, et al. Efficacy of nilotinib (AMN107) in patients (pts) with newly diagnosed, previously untreated Philadelphia chromosome (Ph)-positive chronic myelogenous leukemia in early chronic phase (CML-CP) [abstract]. *Blood*. 2007;110:17a. Abstract 29.
 34. Cortes J, O'Brien S, Jones D, et al. Efficacy of dasatinib in patients (pts) with previously untreated chronic myelogenous leukemia in early chronic phase [abstract]. *Blood*. 2007;110:17a. Abstract 30.
 35. Myers GL, Kimberly MM, Waymack PP, Smith SJ, Cooper GR, Sampson EJ. A reference method laboratory network for cholesterol: a model for standardization and improvement of clinical laboratory measurements. *Clin Chem*. 2000;46:1762-1772.
 36. Baadenhuijsen H, Steigstra H, Cobbaert C, Kuypers A, Weykamp C, Jansen R. Commutability assessment of potential reference materials using a multicenter split-patient-sample between-field-methods (twin-study) design: study within the framework of the Dutch project "Calibration 2000." *Clin Chem*. 2002;48:1520-1525.
 37. Petersen PH, de Verdier C-H, Groth T, Fraser CG, Blaabjerg O, Horder M. The influence of analytical bias on diagnostic misclassifications. *Clinica Chimica Acta*. 1997;260:189-206.
 38. Bennett ST, Connelly DP, Eckfeldt JH. Assessment of split-sample proficiency testing for cholesterol by use of a computer simulation model. *Clin Chem*. 1991;37:497-503.
 39. Branford S, Rudzki Z, Parkinson I, et al. Real-time quantitative PCR analysis can be used as a primary screen to identify patients with CML treated with imatinib who have BCR-ABL kinase domain mutations. *Blood*. 2004;104:2926-2932.
 40. Wang L, Knight K, Lucas C, Clark RE. The role of serial BCR-ABL transcript monitoring in predicting the emergence of BCR-ABL kinase mutations in imatinib-treated patients with chronic myeloid leukemia. *Haematologica*. 2006;91:235-239.
 41. Miller WG. Matrix effects in the measurement and standardization of lipids and lipoproteins. In: Rifai N, Warnick GR, Dominiczak MH, eds. *Handbook of Lipoprotein Testing*. Washington, DC: AACC Press;1997:199-221.
 42. Winn-Deen ES, Helton B, Van Atta R, et al. Development of an integrated assay for detection of BCR-ABL RNA. *Clin Chem*. 2007;53:1593-1600.
 43. Jobbagy Z, van Atta R, Murphy KM, Eshleman JR, Gocke CD. Evaluation of the Cepheid GeneXpert BCR-ABL assay. *J Mol Diagn*. 2007;9:220-227.