Standardisation and consensus guidelines for minimal residual disease assessment in Philadelphia-positive acute lymphoblastic leukemia (Ph+ALL) by real-time quantitative reverse transcriptase PCR of *e1a2 BCR-ABL1*

H Pfeifer[1], G Cazzaniga[2], VHJ van der Velden[3], JM Cayuela[4], B Schäfer[5], O Spinelli[6], S Akiki[7], S Avigad[8], I Bendit[9], K Borg[10], H Cavé[11],L Elia[12], SC Reshmi[13], G Gerrard[14], S Hayette[15], M Hermanson[16], A Juh[17], T Jurcek[18], MC Chillón[19], C Homburg[20], G Martinelli[21], V Kairisto[22], T Lange[23], T Lion[24], MC Mueller[25], F Pane[26], L Rai[27] , C Damm-Welk[28],T Sacha[29], S Schnittger[30], T Touloumenidou[31], H Valerhaugen[32], P Vandenberghe [33], J Zuna[34], H Serve[1], E Herrmann [35], S Markovic[1],[#]JJM van Dongen[3,36] and [#]OG Ottmann[37]

Department of Medicine, Hematology/Oncology, Goethe-University, Frankfurt, Germany[1]
Pediatric Clinic, University of Milan Bicocca, San Gerardo Hospital, Monza, Italy[2]
Department of Immunology, Erasmus MC, University, Medical Center Rotterdam, Rotterdam, The Netherlands[3]
Hematology Laboratory and EA3518, University Hospital Saint-Louis, University Paris Diderot, Paris, France[4]
Kinderspital Zürich, Onkologie Diagnostik Labor, Zürich, Switzerland[5]
Hematology and Bone Marrow Transplant Unit, ASST-Papa Giovanni XXIII, Bergamo, Italy[6]
Regional Genetics Laboratory, Birmingham Women´s NHS Foundation Trust, Birmingham, UK[7]
Pediatric Hematology Oncology, Schneider Children's Medical Center of Israel, Petah Tikva, Israel[8]
Division of Hematology and Hemotherapy, Medical School, University of São Paulo, Brazil[9]
Department of Diagnostic Hematology, Institute of Hematology and Transfusion Medicine, Warsaw, Poland[10]
Department of Genetic Biochemistry, Hopital Robert Debré, Paris, France[11]
Department of Hematology, Laboratory of Molecular Biology, University La Sapienza, Roma, Italy[12]
Institute for Genomic Medicine, Nationwide Children's Hospital and the Ohio State University, Columbus, OH, United States[13]
Imperial Molecular Pathology, Centre for Hematology, Imperial College London, London, UK[14]
Laboratory of Molecular Biology and UMR5239, Centre Hospitalier Lyon-Sud, Hospices Civils de Lyon, Pierre Bénite, Lyon, France[15]
Department of Immunology, Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden[16]
National University Hospital, Division of Pediatric Hematology/Oncology, Department of Pediatrics, Singapore[17]
Department of Internal Medicine-Hematology and Oncology, Masaryk University and University Hospital Brno, Brno, Czech Republic[18]
Department of Hematology, University Hospital of Salamanca (IBSAL) & Cancer Research Centre-IBMCC, (Salamanca University-CSIC), Salamanca, Spain[19]
Sanquin Diagnostics, Department of Experimental Immunhematology/Immuncytology, Amsterdam, The Netherlands[20]
Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy[21]
TYKSLAB, Laboratory of Molecular Genetics, Turku University Hospital, Turku, Finland[22]
Abteilung für Hämatologie und internistische Onkologie, Universität Leipzig, Leipzig, Germany[23]
Childrens Cancer Research Institute, Labdia Labordiagnostik and Medical University, Vienna, Austria[24]
III. Medizinische Klinik, Medizinische Fakultät Mannheim der Universität Heidelberg, Mannheim, Germany[25]
Azienda Ospedaliera Universitaria Federico II di Napoli Università, Naples, Italy[26]
Department of Hematology, Royal Free Campus, University College London, London, United Kingdom[27]
Department of Pediatric Hematology and Oncology, University Hospital, Justus Liebig University Giessen, Giessen, Germany[28]
Hematology Department, Jagiellonain University Krakow, Krakow, Polen[29]
MLL Munich Leukemia Laboratory, Max-Lebsche-Platz 31, Munich, Germany[30]
Hematology Department and HCT Unit, G. Papanicolaou Hospital, Thessaloniki, Greece[31]
Department of Pathology, The Norwegian Radium Hospital, Oslo University Hospital, Oslo; Norway[32]
Center for Human Genetics, University Hospital Leuven and KU Leuven, Leuven, Belgium[33]
CLIP, Department of Paediatric Haematology and Oncology, 2nd Faculty of Medicine, Charles University and UH Motol, Prague, Czech Republic[34]
Institut für Biostatistik und Mathematische Modellierung, Goethe Universität, Frankfurt, Germany[35]
Leiden University Medical Centre, Department of Immunohematology and Blood Transfusion, Leiden, The Netherlands[36]
Cardiff University, School of Medicine, Division of Cancer and Hematology, Cardiff, UK[37]

[#] JJM van Dongen and OG Ottmann contributed equally

Running title: minimal residual disease consensus guidelines in Ph+ALL

Key words: acute lymphoblastic leukemia, BCR-ABL1, minimal residual disease

Address correspondence to:

Heike Pfeifer, M.D.
Department of Medicine, Hematology/Oncology
Goethe University
Theodor-Stern-Kai 7
D-60590 Frankfurt

Tel: ++49 (69) 6301 83044
Fax: ++49 (69) 6301 83046
e-mail: h.pfeifer@em.uni-frankfurt.de

Abstract word count: 199

Text body word count: 4000

Figures: 5

Tables: 2

Suppl. Figures: 7

Abstract

Minimal residual disease (MRD) is a powerful prognostic factor in acute lymphoblastic leukemia (ALL) and is used for patient stratification and treatment decisions, but its precise role in Philadelphia chromosome positive ALL is less clear. This uncertainty results largely from methodological differences relating to the use of real-time quantitative PCR (qRT-PCR) to measure *BCR-ABL1* transcript levels for MRD analysis. We here describe the first results by the EURO-MRD consortium on standardization of qRT-PCR for the *e1a2 BCR-ABL1* transcript in Ph+ALL, designed to overcome the lack of standardisation of laboratory procedures and data interpretation. Standardised use of EAC primer/probe sets and of centrally prepared plasmid standards had the greatest impact on reducing interlaboratory variability. In QC1 the proportion of analyses with *BCR-ABL1/ABL1* ratios within half a log difference were 40/67 (60%) and 52/67 (78%) at 10^{-3} and 36/67 (53%) and 53/67 (79%) at 10^{-4} *BCR-ABL1/ABL1*. Standardized RNA extraction, cDNA synthesis and cycler platforms did not improve results further, whereas stringent application of technical criteria for assay quality and uniform criteria for data interpretation and reporting were essential. We provide detailed laboratory recommendations for the standardized MRD analysis in routine diagnostic settings and in multicenter clinical trials for Ph+ALL.

Introduction

Minimal residual disease (MRD) during treatment for acute lymphoblastic leukemia (ALL) is predictive of outcome and is an accepted measure of response in pediatric and adult ALL.(1-11) The value of MRD for therapeutic decisions (6, 12) relies on the precision and reproducibility of MRD measurements. For Ph-negative ALL, these are typically based on DNA-based quantitative PCR (qPCR) analysis of immunoglobulin (Ig) and T-cell receptor (TCR) gene rearrangements. This methodology has been optimized and standardised internationally as part of the BIOMED concerted action (13, 14) and subsequently the EuroMRD Consortium, formerly European Study Group on MRD detection in ALL (ESG-MRD-ALL). (14-18)

For Ph+ALL, standard MRD monitoring of *BCR-ABL1* by qRT-PCR has been suggested to provide better and earlier prediction of relapse than the Ig/TCR approach.(19) Advantages of reverse transcription real-time quantitative PCR (qRT-PCR) include high sensitivity, speed and low cost, without need for identification of patient-specific MRD markers. Major limitations are the diversity of qRT-PCR protocols, instability and variable expression levels of RNA, risk of cross-contamination and variability in data interpretation and reporting of results. Thus, clinical trials using comparable therapeutic regimens for Ph+ALL have reported widely varying rates and degrees of molecular response after induction and consolidation therapy.(1, 20-26) As precise and reproducible MRD measurements are critical to ensure comparability of results between trials and the validity of MRD-based treatment decisions, standardisation of qRT-PCR procedures, agreement on data interpretation and stringent quality control (QC) measures are essential.(16, 18, 27-30)

Initial efforts to standardize qRT-PCR analysis were conducted within the Europe Against Cancer (EAC) program, which designed and validated EAC primer and

probe sets for the most frequent leukemia-associated fusion genes.(16, 29) In contrast to the chronic myeloid leukemia (CML) community, which initiated international and still ongoing standardisation attempts in 2003,(31, 32) no equivalent optimisation and standardisation has been conducted for Ph+ALL. The importance of high sensitivity and accuracy of quantitating MRD levels in this clinically aggressive disease (33) prompted us to comparatively evaluate the impact of methodological differences in qRT-PCR analysis of *BCR-ABL1*, to assess measures to improve the inter-laboratory variability and reproducibility of *e1a2 BCR-ABL1* detection, and to define guidelines for work-flow and reporting of *e1a2 BCR-ABL1* transcript levels. This study focussed on the *e1a2* breakpoint because of its predominance in Ph+ALL, as opposed to the *e13a2* and *e14a2* breakpoints characteristic of CML.(34) We here report the consensus recommendations of the EuroMRD Consortium for conducting and reporting MRD analyses in patients with Ph+ALL, based on the data generated during the first 10 QC rounds.

Organisation and definition of aims

European multinational cooperative study group trials EsPhALL (35) and EWALL-PH01 (24) for pediatric and adult Ph+ALL respectively, were the first to focus on accuracy and comparability of MRD results generated in the different national reference laboratories. Most of these laboratories participated in standardisation and quality control studies for immunoglobulin and T cell receptor rearrangement-based MRD assessment within the EuroMRD Consortium. Positive experience with these studies led to analogous efforts for the RNA-based quantitation of *BCR-ABL1*. Several laboratories of the European Working Group for Adult ALL (EWALL) as well as those participating in *BCR-ABL1* quantification in CML joined this consortium.

Successive aims and work programs were designed in biannual group meetings, in which results of the studies were discussed. 35 laboratories participated in the first 10 QC rounds, the results of which are provided in this paper. The coordinating centre in Frankfurt was responsible for central preparation and biannual shipment of samples, data collection and analysis.

Materials and Methods

Study Design

Several successive phases of blinded testing were conducted, described for each QC round in supplemental table1. Initially, centrally prepared samples were applied to RNA-extraction, cDNA-synthesis and qRT-PCR for *BCR-ABL1* and *ABL1* with the standard reagents and methods used by the participating laboratories by means of questionnaires. (Suppl. data)

A second phase systematically compared the effect of *i)* RNA-extraction, *ii)* cDNA synthesis, and *iii)* qRT-PCR by applying routine procedures to centrally produced material. cDNA-synthesis and real-time PCR were occasionally repeated on separate days to enhance comparability of the individual experimental steps.

The third phase (QC10) tested the consistency of results and analytical procedures. QC7 was a detailed re-analysis of all data generated in QC rounds 1 to 6. *Q*uantitative measurement of *BCR-ABL1/ABL1* (B/A), and from QC3 onward also *BCR-ABL1/GUS* (B/G), was a common feature of all QC rounds with analyses for *ABL1 or GUS* as control genes (CG), respectively.

Paper tasks

To standardize interpretation of results, additional “paper tasks” were devised in which the laboratories were provided with virtual complex *BCR-ABL1/ABL1* and

BCR-ABL1/GUS measurements and asked for data interpretation and reporting. All results were discussed in depth to clarify controversial and discordant results, providing a well-founded basis for the development of guidelines.

Centralised sample preparation and quality assessment

Lyophilised cell dilutions ranging from 1% to 0.001% *BCR-ABL1* positive SUP-B15 cells in *BCR-ABL1* negative NALM6 cells, RNA and/or cDNA samples were prepared centrally and distributed on dry ice to the participating laboratories. Details of sample preparation and distribution are provided as supplemental data. (Suppl. table 1)

Quality control of the samples prior to shipment was performed centrally by qRT-PCR using the standardized EAC protocols.(25) *BCR-ABL1, ABL1, GUS* were quantitated in either 3 or 2 independent runs, with one exception (QC8). Samples were scored as acceptable when all replicates were within 0.5 log. (Suppl. table 2)

Plasmid standards

Centrally prepared plasmid standards were provided by Martin Müller (pME-3; Medical Faculty Mannheim of the University of Heidelberg, Germany and Nick Cross, Helen White (Wessex Regional GeneticsLaboratory, Salisbury, UK). Both vectors contain a PCR-insert with *e1a2 BCR-ABL1*, *ABL1* and *GUS*, and differ only in the vector itself and the size of the inserts. After large scale production, the plasmid DNA was quantified by spectrophotometry. The plasmid was then linearized and serial dilutions were prepared. They were stabilized in a water-based solution containing 20 ng/µl of msRNA (Roche Diagnostics, Mannheim, Germany). Six dilutions were prepared (2 to $2x10^5$ copies). The corresponding standard curve typically generated a mean slope of -3.45 and an intercept of 41.57 for pME-3 (calculated from QC3, 4,

5, 6). A mean CT value of 27 was generated for 10 000 copies of *ABL1* with a threshold set at 0.1.

Sample analysis by participating laboratories

Participants were asked to analyse the samples using the workflow given in Suppl. table 1. Using questionnaires that were repeatedly sent out throughout the study, laboratories also provided information on their in-house methods and standard approaches to expressing results and determining positivity and negativity.

Statistical Methods

All data were analyzed by the coordinating laboratory in Frankfurt. Results were presented during Euro-MRD meetings in a non-blinded manner. For Box-plot analysis, a 25^{th}–75^{th} percentile box with 10^{th} -90^{th} whiskers was used (median marked by line). Regression analysis was performed with the log B/A or B/G of the sample sets between individual participating laboratories and reference (log(median of B/A or B/G value)). Statistical analysis was performed using the GraphPad Prism™ software package version 6.02 (San Diego, USA).

RESULTS

Trial participants and routine methodology

35 laboratories contributed, 8 from the EWALL, 22 from EuroMRD and 5 additional laboratories with high sample throughput and/or responsibility for CML validation within the EUTOS consortium. The routine methodologies for RNA extraction, cDNA synthesis and qRT-PCR are described under supplemental table 3.

Initial testing phase

Variability of RNA-Extraction

In the first QC, RNA extraction was performed from centrally prepared aliquots of cell dilutions lysed and homogenized in TRIzol® reagent and shipped on dry ice. The amount of RNA extracted from identical aliquots by different laboratories using the same extraction method varied substantially, ranging from 0.66 µg to 49.1 µg (Suppl. Figure 3A).

The *ABL1* copy number following RNA extraction and cDNA preparation likewise showed substantial variability as displayed in the box plots in Supplemental Figure 3B. Uniform use of TRIzol®–based RNA extraction therefore did not improve inter-laboratory variability.

Assay sensitivity in relation to control gene

As a measure of assay sensitivity for the various dilutions of *e1a2 BCR-ABL1* positive cells (total cells per sample: $5x10^6$), the median number of *ABL1* copies generated during the initial testing phase was 5.4×10^4, (range 0 - 2.41×10^6) (n=22 laboratories) (Suppl. Figure 3B). Nine laboratories failed to reach the threshold of 10^4 *ABL1* copies in at least one of seven samples. Overall assay sensitivity across laboratories was thus considered suboptimal, despite sufficient cell numbers available for extraction.

Exemplarily, for *GUS* as CG, the median copy number was 6.4×10^4 (2.13×10^4 - 3.71×10^6) (QC4, using centrally prepared cDNA). With the exception of one outlier, no laboratory measured fewer than 10^4 GUS copies in any sample (Suppl. Figure 4B). The median ABL1 amount also revealed sufficient assay sensitivity in all laboratories (QC4). (Suppl. Figure 4A)

Intra- and inter-laboratory variability was higher when cells instead of cDNA were used as starting material, corresponding to the fewer analytical steps prior to the qRT-PCR reaction. (Figure 2, Suppl. Figure 1 & 2).

Concordance of BCR-ABL1 transcript levels between laboratories

As shown in figure 1, the *BCR-ABL1/ABL1* ratios varied. Across all QC rounds and all samples, the inter-laboratory variability of *BCR-ABL1* transcripts levels correlated inversely with the number of transcripts analyzed (Figure 1). For example, in QC1, the range of B/A ratios in samples with median B/A ratios of 1.99×10^{-1}, 4.41×10^{-2}; 2.53×10^{-2} covered a range of 2.26 log, 1.82 log and 1.81 log, respectively. At these dilutions, 61/67 (91%), 62/67 (93%) and 56/67 (83%) of analyses measuring *BCR-ABL1/ABL1* ratio were within 1 log of the median and there were no false negatives. (Figure 1A)

For the median B/A of 2.5 x 10-3, 40/67 (60%) analysis were within 0.25 log, 52 of 67 (78%) analysis were within 0.5 log, 62 of 67 (93%) analyses were within 1 log above or below the median. The overall range of B/A ratios was substantially broader (0 to 0.03) with 2 of 22 laboratories measuring no *BCR-ABL1* in this sample. (Figure 1A)

Deviation from the median at the lowest *BCR-ABL1* transcript numbers (sample 6 in Figure 1A; median 2.26×10^{-5}) was large, ranging from negative to 2.09 log above the median. In 21 of 67 analyses (31%) *e1a2 BCR-ABL1* was negative in at least one replicate. 8/22 (36%) laboratories failed to measure this sample as positive.

Data obtained with *GUS* instead of *ABL1* as CG were comparable (Figure 1C & 1D). The substantially greater variability of B/A or B/G ratios and a higher proportion of negative results at low *BCR-ABL1* transcript levels are attributable to low transcript numbers often being below the quantitative range of the qRT-PCR assay.

To determine linearity and sensitivity of the results from each participating laboratory, we performed a linear regression analysis in which we compared the log B/A ratios obtained by the reference (the median of B/A ratios) and participating laboratories. As an example, twenty of 22 (90.9%) local laboratories obtained linear results (R^2 >0.98) when plotted on a log scale (QC1, Suppl. Figure 5A). The non-linear fit of data from 2 laboratories suggested an unsatisfactory amplification efficiency of the respective PCR technique within the meaningful diagnostic range. Results with GUS as CG were similar (QC4, Suppl. Figure 5B).

Impact of methodology on inter-laboratory variability

Poor reproducibility of *BCR-ABL1* quantitation below quantitative range (Figure 1, Suppl. Figure 1 & 2) prompted us to assess whether cDNA synthesis performed centrally versus locally improved inter-laboratory variability. Centralization RNA extraction and cDNA synthesis resulted in only modest improvement of the inter-laboratory variability of B/A ratios. As shown in Figure 2, the interquartile range was slightly smaller when cDNA was produced centrally instead of locally ($2.07x10^{-3}$ - $4.78x10^{-3}$ *vs.* $2.63x10^{-3}$ - $4.10x10^{-3}$; QC6, sample 3). Similarly, the frequency of false-negative results obtained with centrally (4/26; 15%) or locally (2/26; 8%) synthesised cDNA was similar at very low B/A ratios, e.g. 3.41×10^{-5} (Fig. 2, QC6, sample 6). Thus, centralisation of cDNA synthesis did not significantly improve inter-laboratory variability when transcript levels were below the quantitative range (10 copies). SuperScript and M-MLV, the principal enzymes used for cDNA synthesis, were comparable in terms of the B/A ratios (Figure 3A) and the *ABL1* amount (Suppl. Figure 6A), irrespective of the *BCR-ABL1* transcript level. Comparison of the cycler types revealed no difference with respect to either the B/A ratio (Figure 3B) or *ABL1* amount (Suppl. Figure 6B).

Inter-laboratory variability decreased with the use of a centrally prepared plasmid standards (Fig. 4, Suppl. Fig. 1& 2, 7A & B and Supplemental data).

Sensitivity and specificity

The frequency of false negative and false-positive results was analysed in all QC rounds by evaluating blinded negative and low-positive samples (Table 1). False-positive results were generated by 12% to 27% of laboratories in 6/10 QC rounds. No individual laboratory was disproportionally represented. The frequency of false negative results differed by transcript level: at B/A ratios between $1x10^{-4}$ and $5x10^{-4}$, any false negativity occurred in only 2/9 QC rounds, whereas false negative results occurred in 6 of 6 QC rounds at B/A ratios below 10^{-4} (Table 1).

Guidelines for interpretation

Definition of quantitative range and sensitivity

The paper tasks revealed that laboratories initially used different criteria for reporting qRT-PCR results, particularly with low-level positive samples, e.g. when single replicates were positive below 10 *BCR-ABL1* copies. We therefore developed guidelines for data interpretation analogous to those described for qPCR of Ig/TCR rearrangements (17) (Table 2).

The main difference between these guidelines is a more stringent definition for the slope (3.2-3.7) compared with the guidelines published for Ig/TCR quantitation (3.1-3.9). The BCR-ABL1 assay is a uniformassay and therefore likely to be equally efficient between differentpatients, whereas IG/TCR assays are patient-specific and therefore highly variable, which accounts for the higher variability in the efficacy of the PCR reaction and consequently in a wider range for the slope of the standard curve. Based on evaluation of commercial plasmid standard curves (n=65, minor-

BCR-ABL1 standards, QC 1-6), 50% of the slopes were between -3.46 and -3.386 and 61/65 (94%) were between -3.7 and -3.2, which therefore was chosen pragmatically as acceptable range. 154 data sets from in-house plasmid standard curves (QC 1-6) were evaluated for defining the acceptable Delta-CT: 2.6 – 4.0 (25^{th} -75^{th} percentile 2.882 - 3.633, range 1.4 – 6.4) (Figure 5).

Discussion

MRD is an independent prognostic factor for Ph+ALL and may be used as a trigger for therapeutic intervention(21-23, 36-42) but validity and comparability of MRD results obtained by *BCR-ABL1* transcript analysis are uncertain given the lack of standardisation. Our initial comparison of laboratory procedures among participating laboratories revealed almost universal use of EAC primer/probe sets (16), which were therefore accepted as existing standard. Otherwise, no two laboratories initially followed exactly the same protocols, prompting us to evaluate the impact of other workflow conditions in successive QC rounds. Data obtained during the consecutive QA rounds showed that assay sensitivity was particularly related to an individual laboratory´s experience with a qRT-PCR protocol rather than to a specific method. Likewise, *BCR-ABL1* and control gene copy numbers were not systematically affected by the choice of enzymes or type of thermocycler. Successive QC rounds showed that adopting the EAC primer/probe sets, common plasmid standards, the use of defined technical criteria for quality assurance, and common definitions of molecular response and data interpretation, allowed accurate measurement of MRD in Ph+ALL across numerous laboratories. There was no apparent benefit in more

comprehensive standardisation of other components of the laboratory workflow. However, use of a secondary reference panel for *BCR-ABL1* quantification, as recently published for MRD analysis in CML(43) may further improve the accuracy and consistency of MRD results across laboratories.

The choice of control gene presently remains controversial. *ABL1* is used most commonly and yielded results similar to those obtained with GUS, if MRD levels were low. However, at higher *BCR-ABL1* levels, such as at diagnosis, the use of ABL as control gene will underestimate *BCR-ABL1* levels relative to control gene levels (since the EAC *ABL*1 PCR will also detect *BCR-ABL1* transcripts, control gene levels will be over-estimated at high MRD levels). Since *BCR-ABL1* MRD levels in ALL are frequently expressed relative to levels at diagnosis, the use of GUS as control gene seems more appropriate and is therefore strongly recommended.

Throughout the QC rounds, the distinction between detectable and undetectable disease was a critical issue at ≤10 *BCR-ABL1* copies. It became apparent that this challenge was initially compounded by differences in the way individual laboratories interpreted data. By conducting paper tasks, it was demonstrated that variable approaches to interpretation of the same data sets led to substantial discrepancies in the reporting of results, comparable to our experiences for antigen receptor gene rearrangements-based MRD analysis.(17) This problem was largely alleviated by formulating guidelines for data interpretation and reporting that were adhered to by all consortium members, and were reinforced by regular discussions during meetings of the consortium. Despite the introduced guidelines, classification of negative samples as questionably positive still occurred in all QC rounds; this was however

not attributable to specific sites, indicating there was no systematic analytical flaw affecting individual laboratories.

Although transcript values fell within a half log range in 60% of replicate analyses and within 1 log in 78% of replicate analyses (QC1) at a transcript level of 10^{-3}, the total range of values spanned a range of up to 3 log. Variability was even more pronounced at transcript levels of $\leq 10^{-4}$ and coincided with a clinically relevant rate of false positive and false negative results. As expected from comprehensive standardisation efforts conducted in the CML field, use of common plasmid standards instead of cell lines or e1a2 oligonucleotides in the workflow of all laboratories had a profoundly positive impact on interlaboratory variability. Results obtained with commercially available, centrally prepared or in-house plasmids were comparable. The importance of plasmid calibrators to accurately determine the number of *BCR-ABL1* and control gene transcripts in CML has led to the development of an internationally accepted certified plasmid reference material for *BCR-ABL1* in CML.(44) A similar approach would be desirable for *e1a2 BCR-ABL1* to ensure international comparability of MRD results for Ph+ALL. Meanwhile, laboratories performing MRD testing for multicentre clinical trials of Ph+ALL should use identical, centrally prepared or commercial plasmid reference material.

In CML, a test-specific IS conversion factor to enable comparability of *BCR-ABL1* assays was developed and successfully implemented in a large number of laboratories, but is expensive and time consuming.(45, 46) Our consensus to not implement such a conversion factor for MRD analysis in Ph+ALL reflects reliance on a far smaller number of laboratories that also usually function as national or cooperative group reference laboratories. Moreover, patients with Ph+ALL are

treated predominantly within national cooperative group trials rather than in a community setting. This different environment with a limited number of laboratories facilitates regular participation in QC rounds that are necessary to reinforce adherence to agreed guidelines, facilitate data consistency and ensure comparability of MRD results, as shown previously for the optimization of MRD analysis based on detection of Ig/TCR rearrangements.(14) The EURO-MRD now conducts QC rounds for *e1a2 BCR-ABL1* analysis twice per year.

As clinical management of Ph+ALL is far more heterogeneous than that of CML, depend on diverse treatment protocols and the clinical context, e.g. adult or pediatric or pre-versus post-transplant settings, it was not our aim to define specific clinical recommendations for MRD testing in Ph+ALL. However, our guidelines listed in Table 2 together with the consensus terminology of the second international symposium on MRD assessment (15) provide a technical basis for uniform and correct interpretation of MRD results. Nevertheless, a number of clinical recommendations can be made based on discussions within our consortium and published data.

First, different thresholds and time points have been used by different national cooperative study groups to define molecular remission for Ph+ALL. Similarly, timing of MRD testing is important but it´s predictive value depends considerably on the specific treatment regimen employed, although several underlying principles apply. In general, a good early response after induction therapy is more informative of a favourable outcome, whereas persisting levels after consolidation are indicative of a poor prognosis.(47) During the early treatment period, MRD testing should be frequent, e.g. after every other cycle of chemotherapy, with more frequent analyses

in case of high or rising *BCR-ABL1* transcript levels. Testing at a single time-point may be informative but serial testing is recommended as it provides important information on disease dynamics.(48) In patients undergoing SCT, MRD should be determined within a short time window prior to start of conditioning, as the MRD level before SCT may be predictive of outcome.(49, 50) The first post-transplant MRD test should be performed approximately 4 weeks after SCT and be followed by more frequent analyses than is customary in CML. The EBMT recommends serial assessment of BM every 6 weeks and PB every 3 weeks, based on a prospective randomized trial investigating two schedules of TKI maintenance after SCT.(37) It should be stressed that our analyses was not designed to assign clinical relevance to a specific MRD level but to provide a framework for standardised MRD analysis, allowing comparison of MRD data from multiple future studies and thereby facilitating identification of most informative time points and cut-off levels for Ph+ALL patients.

Second, the material used for MRD analysis clearly impacts on results. As for other BCP-ALL, and in contrast to CML, bone marrow is the preferred source for quantification of MRD in Ph+ALL, since MRD levels typically are 1 – 3 log higher than in peripheral blood, without good correlation between BM and PB.(51)

In occasional patients it may be difficult to obtain good BM aspirates, particularly after SCT with TBI-based conditioning, in which case MRD sensitivity may be higher in PB. A large proportion of BM tests carried out prior to extramedullary relapse may be negative;(52-54) definite MRD positivity in PB in conjunction with negative results in bone marrow may be associated with invasion of extramedullary sites.(55) It is therefore advisable to send BM and PB samples in parallel.

Third, a good quality bone marrow sample, i.e. the first aspirate, should be used for MRD analysis. The aspirate should be collected after hematopoietic recovery

following chemotherapy to ensure sufficient cellularity and assay sensitivity. Prompt delivery to the laboratory conducting the MRD analysis is critical, shipment delays over 48 h need to be avoided since this may impact transcript levels.(56)

The upcoming QC and standardisation efforts will be extended to include quantification of Major-*BCR-ABL1* transcripts, as this breakpoint is present in approximately 30% of patients with Ph+ALL. Reagents such as secondary *BCR-ABL1* reference panels developed by the CML field should prove instrumental in more rapidly reaching consensus on quantification of this second largest type of *BCR-ABL1* transcripts.

In conclusion, our guidelines provide a robust framework for precise and reproducible qRT-PCR based analyses of *e1a2 BCR-ABL1* transcripts and will facilitate valid comparison of MRD results between clinical trials for Ph+ALL. Adhering to this standard set of recommendations will help answering clinically unresolved questions about the prognostic and predictive value of specific *BCR-ABL1* thresholds and dynamics in distinct therapeutic settings, and help resolve a number of clinical management issues that depend on accurate quantification of *BCR-ABL1* transcripts.

Authorship

Contribution: H.P, O.G.O., G.C., VHJvdV, JMC, JJMvD , O.O. designed the study, analyzed data, wrote the manuscript; H.P. ,G.C., V.H.J.vdV., J.M.C., B.S., O.S., S.A., S.A., I.B., K.B., H.C., L.E., J. G.-F., G.G., S.H., M.H., A.J., T.J., M.C.C., C.H., I.I., V.K., T.L., T.L., M.C.M., F.P., L.R., C.DW., T.S., S.S., T.T., H.V., P.VdB., J.Z., H.S.,

S.M. performed the molecular analyses, collected data, H.P., E.H. performed the statistical analysis.

Conflict of Interest

VHJvdV: contract research for Amgen, Roche, Pfizer, Janssen and BD Biosciences; consultancy fees from Celgene. JJMvD: contract research for Roche, Amgen and BD Biosciences. HP: research support from Ipsogen. The other authors declare no conflict of interest.

Acknowledgements

We are grateful to the teams of technicians, particularly Brigitte Gehrke, Doreen Badowski, Gabriele Lippok from the coordinating lab for their excellent work. This study was supported financially by the Deutsche José Carreras Leukämie-Stiftung e.V. (Förderkennzeichen: R09/11 and R12/01). OGO held a Stiftungsprofessur für Molekulare Therapieforschung der Deutschen José Carreras Leukämie Stiftung.

Figure legends

Figure 1: Inter-laboratory variability of *e1a2 BCR-ABL1* transcript levels using *ABL1* or *GUS* as CGs using in-house methods.

BCR-ABL1 quantitation is represented by the ratio of *BCR-ABL1* copy number and *ABL1* copy number plotted on the y-axis (log-transformed). On the x-axis, the sample ID represents different dilutions of centrally prepared TRIzol®-stabilized cells (panel A), lyophilised cells (panel B & D) or cDNA (panel C) that were measured by the participating laboratories in QC rounds 1 and 10 for *ABL1* and QC rounds 4 and 10 for *GUS*. In case of repeat analyses, all data sets are integrated. The median corresponds to the horizontal line, the box refers to the range defined by the 25th and 75th percentiles, the whiskers refer to the range defined by the 10 - 90 percentiles. Outliers are indicated by dots. Each datapoint represents the mean of triplicate analyses of each measurement performed by the participating centers.

Figure 2: Inter-laboratory variability of *e1a2 BCR-ABL1* quantitation in relation to centrally versus locally prepared cDNA as starting material prior to qRT-PCR

In QC6, a pool of centrally prepared RNA was separated into aliquots which were either sent out to the individual laboratories (n=26) who then performed the cDNA synthesis followed by qRT-PCR, or were processed to cDNA, again pooled and then sent out by the central laboratory. With these samples, the participating laboratories performed only the quantitative PCR. The ratio is defined as the ratio between the copy numbers of *BCR-ABL1* and *ABL1* plotted on the y-axis (log-transformed). The median corresponds to the horizontal line, the box refers to the range defined by the 25th and 75th percentiles, the whiskers refer to the range defined by the 10-90 percentiles. Outliers are indicated by dots. Each datapoint represents the mean of triplicate analyses of each measurement performed by the participating centers.

Figure 3: Inter-laboratory variability of e1a2 *BCR-ABL1/ABL1* quantitation in relation to the method of cDNA synthesis and cycler platform

Samples were supplied as centrally produced RNA or cDNA, and the participating laboratories were asked to perform either cDNA-synthesis and real-time PCR or real-time PCR alone (QC6, Sample 1 & 5). On the y-axis the *BCR-ABL1/ABL1* ratio of all results coming from the individual labs are shown (log-transformed), the LAB ID is given on the x-axis. Each dot corresponds to the median of three replicates. Panel A is a comparison of different cDNA enzymes using MMLV or Superscript. Panel B is a comparison of different cycler platforms.Each datapoint represents the mean of triplicate analyses of each measurement performed by the participating centers.

Figure 4: Inter-laboratory variability of e1a2 *BCR-ABL1* transcript quantitation using identical plasmid standards.

In QC10, samples with an intermediate (*BCR-ABL1/ABL1* ratio 10^{-3}) *e1a2 BCR-ABL1* transcript level were provided as cell pellets. Laboratories were asked to perform RNA extraction, followed by two rounds of cDNA synthesis and qRT-PCR performed on separate days. The *e1a2 BCR-ABL1/ABL1* ratios for these paired samples are depicted as a quadrat for the first and a triangle for the second cDNA synthesis for each individual laboratory, as indicated by the lab ID on the x axis. On the right hand side of the two panels, the horizontal line represents the median of all measurements performed by all laboratories, the box refers to the range defined by the 25^{th} and 75^{th} percentiles, the whiskers refer to the range defined by the 10- 90 percentiles. Outliers are indicated by dots. Each datapoint represents the mean of triplicate analyses of each measurement performed by the participating centers.

The y axis is scaled to cover 2 logs (log scale).

Figure 5: Ability of laboratories to detect *BCR-ABL1* copies within the QR

An example representing one of 156 datasets showing the ability of participating laboratories to detect 10 to 100 *BCR-ABL1* copies from the plasmid standard (using in-house plasmids) within the quantitative range defined in QC6. On the y-axis, the delta-CT between 10 and 100 copies is shown. The lab ID is given on the x-axis.

On the right hand side, the median corresponds to the horizontal line, the box refers to the range defined by the 25th and 75th percentiles, the whiskers refer to the range defined by the minimum and maximum. Abbreviations: QR, quantitative range

References

1. Bassan R, Spinelli O, Oldani E, Intermesoli T, Tosi M, Peruta B, et al. Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL). Blood. 2009;113(18):4153-62.
2. Brisco J, Hughes E, Neoh SH, Sykes PJ, Bradstock K, Enno A, et al. Relationship between minimal residual disease and outcome in adult acute lymphoblastic leukemia. Blood. 1996;87(12):5251-6.
3. Bruggemann M, Raff T, Flohr T, Gokbuget N, Nakao M, Droese J, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. Blood. 2006;107(3):1116-23.
4. Cave H, van der Werff ten Bosch J, Suciu S, Guidal C, Waterkeyn C, Otten J, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer--Childhood Leukemia Cooperative Group. N Engl J Med. 1998;339(9):591-8.
5. Conter V, Bartram CR, Valsecchi MG, Schrauder A, Panzer-Grumayer R, Moricke A, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood. 2010;115(16):3206-14.
6. Gokbuget N, Kneba M, Raff T, Trautmann H, Bartram CR, Arnold R, et al. Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. Blood. 2012;120(9):1868-76.
7. Mortuza FY, Papaioannou M, Moreira IM, Coyle LA, Gameiro P, Gandini D, et al. Minimal residual disease tests provide an independent predictor of clinical outcome in adult acute lymphoblastic leukemia. J Clin Oncol. 2002;20(4):1094-104.
8. van der Velden VH, Joosten SA, Willemse MJ, van Wering ER, Lankester AW, van Dongen JJ, et al. Real-time quantitative PCR for detection of minimal residual disease before allogeneic stem cell transplantation predicts outcome in children with acute lymphoblastic leukemia. Leukemia. 2001;15(9):1485-7.
9. Borowitz MJ, Pullen DJ, Shuster JJ, Viswanatha D, Montgomery K, Willman CL, et al. Minimal residual disease detection in childhood precursor-B-cell acute lymphoblastic leukemia: relation to other risk factors. A Children's Oncology Group study. Leukemia. 2003;17(8):1566-72.
10. Conter V, Valsecchi MG, Parasole R, Putti MC, Locatelli F, Barisone E, et al. Childhood high-risk acute lymphoblastic leukemia in first remission: results after chemotherapy or transplant from the AIEOP ALL 2000 study. Blood. 2014;123(10):1470-8.
11. Bader P, Hancock J, Kreyenberg H, Goulden NJ, Niethammer D, Oakhill A, et al. Minimal residual disease (MRD) status prior to allogeneic stem cell transplantation is a powerful predictor for post-transplant outcome in children with ALL. Leukemia. 2002;16(9):1668-72.
12. Topp MS, Kufer P, Gokbuget N, Goebeler M, Klinger M, Neumann S, et al. Targeted therapy with the T-cell-engaging antibody blinatumomab of chemotherapy-refractory minimal residual disease in B-lineage acute lymphoblastic leukemia patients results in high response rate and prolonged leukemia-free survival. J Clin Oncol. 2011;29(18):2493-8.
13. Pongers-Willemse MJ, Seriu T, Stolz F, d'Aniello E, Gameiro P, Pisa P, et al. Primers and protocols for standardized detection of minimal residual disease in acute lymphoblastic leukemia using immunoglobulin and T cell receptor gene rearrangements and TAL1 deletions as PCR targets: report of the BIOMED-1 CONCERTED ACTION: investigation of minimal residual disease in acute leukemia. Leukemia. 1999;13(1):110-8.
14. van der Velden VH, Panzer-Grumayer ER, Cazzaniga G, Flohr T, Sutton R, Schrauder A, et al. Optimization of PCR-based minimal residual disease diagnostics for childhood acute lymphoblastic leukemia in a multi-center setting. Leukemia. 2007;21(4):706-13.

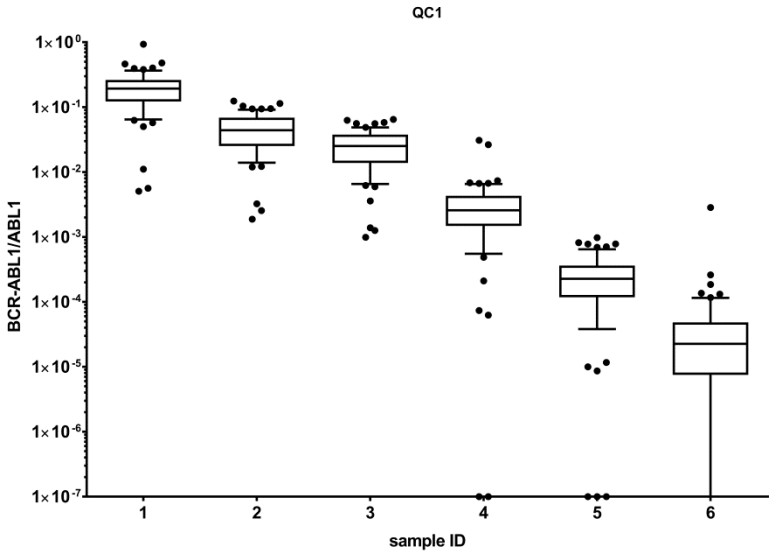
15. Bruggemann M, Schrauder A, Raff T, Pfeifer H, Dworzak M, Ottmann OG, et al. Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. Leukemia. 2010;24(3):521-35.
16. Gabert J, Beillard E, van der Velden VH, Bi W, Grimwade D, Pallisgaard N, 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(12):2318-57.
17. van der Velden VH, Cazzaniga G, Schrauder A, Hancock J, Bader P, Panzer-Grumayer ER, et al. Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. Leukemia. 2007;21(4):604-11.
18. van der Velden VH, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. Leukemia. 2003;17(6):1013-34.
19. van Dongen JJ, van der Velden VH, Bruggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. Blood. 2015;125(26):3996-4009.
20. Chalandon Y, Thomas X, Hayette S, Cayuela JM, Abbal C, Huguet F, et al. Randomized study of reduced-intensity chemotherapy combined with imatinib in adults with Ph-positive acute lymphoblastic leukemia. Blood. 2015;125(24):3711-9.
21. Foa R, Vitale A, Vignetti M, Meloni G, Guarini A, De Propris MS, et al. Dasatinib as first-line treatment for adult patients with Philadelphia chromosome-positive acute lymphoblastic leukemia. Blood. 2011;118(25):6521-8.
22. Lee S, Kim DW, Cho BS, Yoon JH, Shin SH, Yahng SA, et al. Impact of minimal residual disease kinetics during imatinib-based treatment on transplantation outcome in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia. 2012;26(11):2367-74.
23. Ravandi F, Jorgensen JL, Thomas DA, O'Brien S, Garris R, Faderl S, et al. Detection of MRD may predict the outcome of patients with Philadelphia chromosome-positive ALL treated with tyrosine kinase inhibitors plus chemotherapy. Blood. 2013;122(7):1214-21.
24. Rousselot P, Coude MM, Gokbuget N, Gambacorti Passerini C, Hayette S, Cayuela JM, et al. Dasatinib and low-intensity chemotherapy in elderly patients with Philadelphia chromosome-positive ALL. Blood. 2016;128(6):774-82.
25. Thomas DA, Faderl S, Cortes J, O'Brien S, Giles FJ, Kornblau SM, et al. Treatment of Philadelphia chromosome-positive acute lymphocytic leukemia with hyper-CVAD and imatinib mesylate. Blood. 2004;103(12):4396-407.
26. Wassmann B, Pfeifer H, Goekbuget N, Beelen DW, Beck J, Stelljes M, et al. Alternating versus concurrent schedules of imatinib and chemotherapy as front-line therapy for Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL). Blood. 2006;108(5):1469-77.
27. Schrappe M. Minimal residual disease: optimal methods, timing, and clinical relevance for an individual patient. Hematology Am Soc Hematol Educ Program. 2012;2012:137-42.
28. van der Velden VH, Boeckx N, van Wering ER, van Dongen JJ. Detection of minimal residual disease in acute leukemia. J Biol Regul Homeost Agents. 2004;18(2):146-54.
29. van Dongen JJ, Macintyre EA, Gabert JA, Delabesse E, Rossi V, Saglio G, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. Leukemia. 1999;13(12):1901-28.
30. Beillard E, Pallisgaard N, van der Velden VH, Bi W, Dee R, van der Schoot E, 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(12):2474-86.
31. Branford S, Cross NC, Hochhaus A, Radich J, Saglio G, Kaeda J, et al. Rationale for the recommendations for harmonizing current methodology for detecting BCR-ABL transcripts in patients with chronic myeloid leukaemia. Leukemia. 2006;20(11):1925-30.

32. Hughes T, Deininger M, Hochhaus A, Branford S, Radich J, Kaeda J, 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(1):28-37.
33. Ottmann OG, Pfeifer H. Management of Philadelphia chromosome-positive acute lymphoblastic leukemia (Ph+ ALL). Hematology Am Soc Hematol Educ Program. 2009:371-81.
34. De Klein A, Hagemeijer A, Bartram CR, Houwen R, Hoefsloot L, Carbonell F, et al. bcr rearrangement and translocation of the c-abl oncogene in Philadelphia positive acute lymphoblastic leukemia. Blood. 1986;68(6):1369-75.
35. Biondi A, Schrappe M, De Lorenzo P, Castor A, Lucchini G, Gandemer V, et al. Imatinib after induction for treatment of children and adolescents with Philadelphia-chromosome-positive acute lymphoblastic leukaemia (EsPhALL): a randomised, open-label, intergroup study. Lancet Oncol. 2012;13(9):936-45.
36. Lee S, Kim YJ, Chung NG, Lim J, Lee DG, Kim HJ, et al. The extent of minimal residual disease reduction after the first 4-week imatinib therapy determines outcome of allogeneic stem cell transplantation in adults with Philadelphia chromosome-positive acute lymphoblastic leukemia. Cancer. 2009;115(3):561-70.
37. Pfeifer H, Wassmann B, Bethge W, Dengler J, Bornhauser M, Stadler M, et al. Randomized comparison of prophylactic and minimal residual disease-triggered imatinib after allogeneic stem cell transplantation for BCR-ABL1-positive acute lymphoblastic leukemia. Leukemia. 2013;27(6):1254-62.
38. Wassmann B, Pfeifer H, Stadler M, Bornhauser M, Bug G, Scheuring UJ, et al. Early molecular response to posttransplantation imatinib determines outcome in MRD+ Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL). Blood. 2005;106(2):458-63.
39. Yanada M, Sugiura I, Takeuchi J, Akiyama H, Maruta A, Ueda Y, et al. Prospective monitoring of BCR-ABL1 transcript levels in patients with Philadelphia chromosome-positive acute lymphoblastic leukaemia undergoing imatinib-combined chemotherapy. Br J Haematol. 2008;143(4):503-10.
40. Bassan R, Rossi G, Pogliani EM, Di Bona E, Angelucci E, Cavattoni I, et al. Chemotherapy-phased imatinib pulses improve long-term outcome of adult patients with Philadelphia chromosome-positive acute lymphoblastic leukemia: Northern Italy Leukemia Group protocol 09/00. J Clin Oncol. 2010;28(22):3644-52.
41. Radich J, Gehly G, Lee A, Avery R, Bryant E, Edmands S, et al. Detection of bcr-abl transcripts in Philadelphia chromosome-positive acute lymphoblastic leukemia after marrow transplantation. Blood. 1997;89(7):2602-9.
42. Zaliova M, Fronkova E, Krejcikova K, Muzikova K, Mejstrikova E, Stary J, et al. Quantification of fusion transcript reveals a subgroup with distinct biological properties and predicts relapse in BCR/ABL-positive ALL: implications for residual disease monitoring. Leukemia. 2009;23(5):944-51.
43. Cross NC, White HE, Ernst T, Welden L, Dietz C, Saglio G, et al. Development and evaluation of a secondary reference panel for BCR-ABL1 quantification on the International Scale. Leukemia. 2016;30(9):1844-52.
44. White H, Deprez L, Corbisier P, Hall V, Lin F, Mazoua S, et al. A certified plasmid reference material for the standardisation of BCR-ABL1 mRNA quantification by real-time quantitative PCR. Leukemia. 2015;29(2):369-76.
45. Branford S, Fletcher L, Cross NC, Muller MC, Hochhaus A, Kim DW, et al. 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. Blood. 2008;112(8):3330-8.
46. Cross NC, Hochhaus A, Muller MC. Molecular monitoring of chronic myeloid leukemia: principles and interlaboratory standardization. Ann Hematol. 2015;94 Suppl 2:S219-25.
47. Bruggemann M, Kotrova M. Minimal residual disease in adult ALL: technical aspects and implications for correct clinical interpretation. Blood Adv. 2017;1(25):2456-66.
48. Scheuring UJ, Pfeifer H, Wassmann B, Bruck P, Gehrke B, Petershofen EK, et al. Serial minimal residual disease (MRD) analysis as a predictor of response duration in Philadelphia-positive acute lymphoblastic leukemia (Ph+ALL) during imatinib treatment. Leukemia. 2003;17(9):1700-6.

49. Lussana F, Intermesoli T, Gianni F, Boschini C, Masciulli A, Spinelli O, et al. Achieving Molecular Remission before Allogeneic Stem Cell Transplantation in Adult Patients with Philadelphia Chromosome-Positive Acute Lymphoblastic Leukemia: Impact on Relapse and Long-Term Outcome. Biol Blood Marrow Transplant. 2016;22(11):1983-7.
50. Nishiwaki S, Imai K, Mizuta S, Kanamori H, Ohashi K, Fukuda T, et al. Impact of MRD and TKI on allogeneic hematopoietic cell transplantation for Ph+ALL: a study from the adult ALL WG of the JSHCT. Bone Marrow Transplant. 2016;51(1):43-50.
51. van der Velden VH, Jacobs DC, Wijkhuijs AJ, Comans-Bitter WM, Willemse MJ, Hahlen K, et al. Minimal residual disease levels in bone marrow and peripheral blood are comparable in children with T cell acute lymphoblastic leukemia (ALL), but not in precursor-B-ALL. Leukemia. 2002;16(8):1432-6.
52. Cave H, Guidal C, Rohrlich P, Delfau MH, Broyart A, Lescoeur B, et al. Prospective monitoring and quantitation of residual blasts in childhood acute lymphoblastic leukemia by polymerase chain reaction study of delta and gamma T-cell receptor genes. Blood. 1994;83(7):1892-902.
53. Seriu T, Yokota S, Nakao M, Misawa S, Takaue Y, Koizumi S, et al. Prospective monitoring of minimal residual disease during the course of chemotherapy in patients with acute lymphoblastic leukemia, and detection of contaminating tumor cells in peripheral blood stem cells for autotransplantation. Leukemia. 1995;9(4):615-23.
54. van Dongen JJ, Seriu T, Panzer-Grumayer ER, Biondi A, Pongers-Willemse MJ, Corral L, et al. Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. Lancet. 1998;352(9142):1731-8.
55. Coustan-Smith E, Sancho J, Hancock ML, Razzouk BI, Ribeiro RC, Rivera GK, et al. Use of peripheral blood instead of bone marrow to monitor residual disease in children with acute lymphoblastic leukemia. Blood. 2002;100(7):2399-402.
56. van der Velden VH, Boeckx N, Gonzalez M, Malec M, Barbany G, Lion T, et al. Differential stability of control gene and fusion gene transcripts over time may hamper accurate quantification of minimal residual disease--a study within the Europe Against Cancer Program. Leukemia. 2004;18(4):884-6.

Figure 1: Inter-laboratory variability of *e1a2 BCR-ABL1* transcript levels using *ABL1* or *GUS* as CGs using in-house methods

A: *e1a2 BCR-ABL1/ABL1* (in-house QC1)



B: *e1a2 BCR-ABL1/ABL1* (in-house QC10)

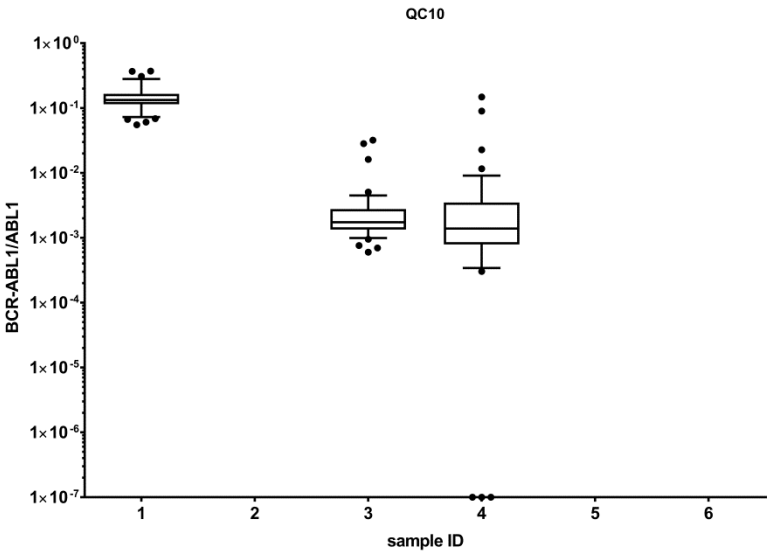
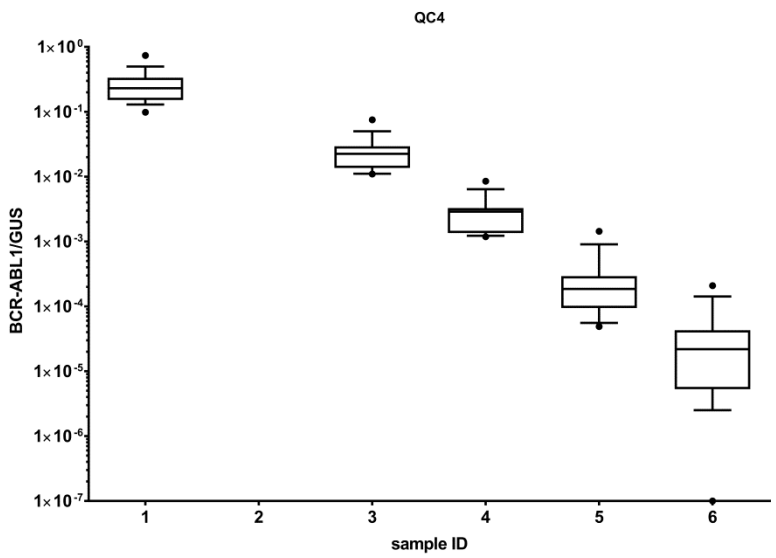


Figure 1: Inter-laboratory variability of *e1a2 BCR-ABL1* transcript levels using *ABL1* or *GUS* as CGs using in-house methods

C: BCR-ABL1/GUS (in-house QC4)



D: BCR-ABL1/GUS (in-house QC10)

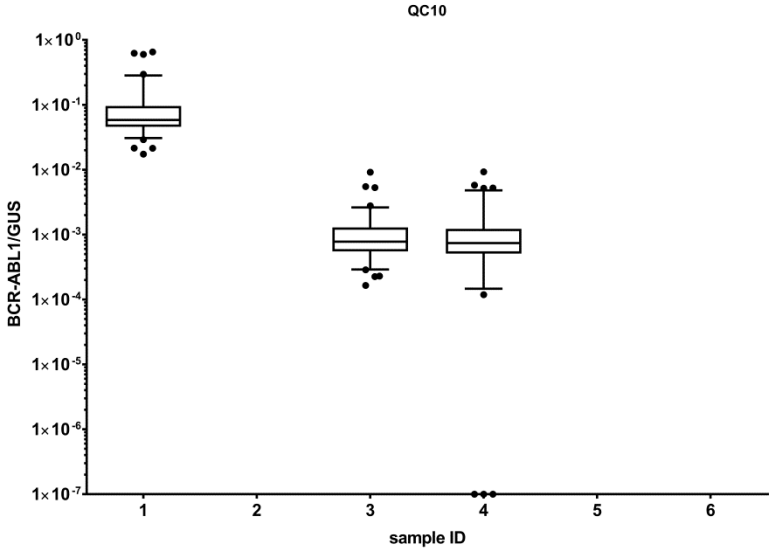


Figure 2: Inter-laboratory variability of *BCR-ABL1* quantitation in relation to centrally versus locally prepared cDNA as starting material prior to qRT-PCR

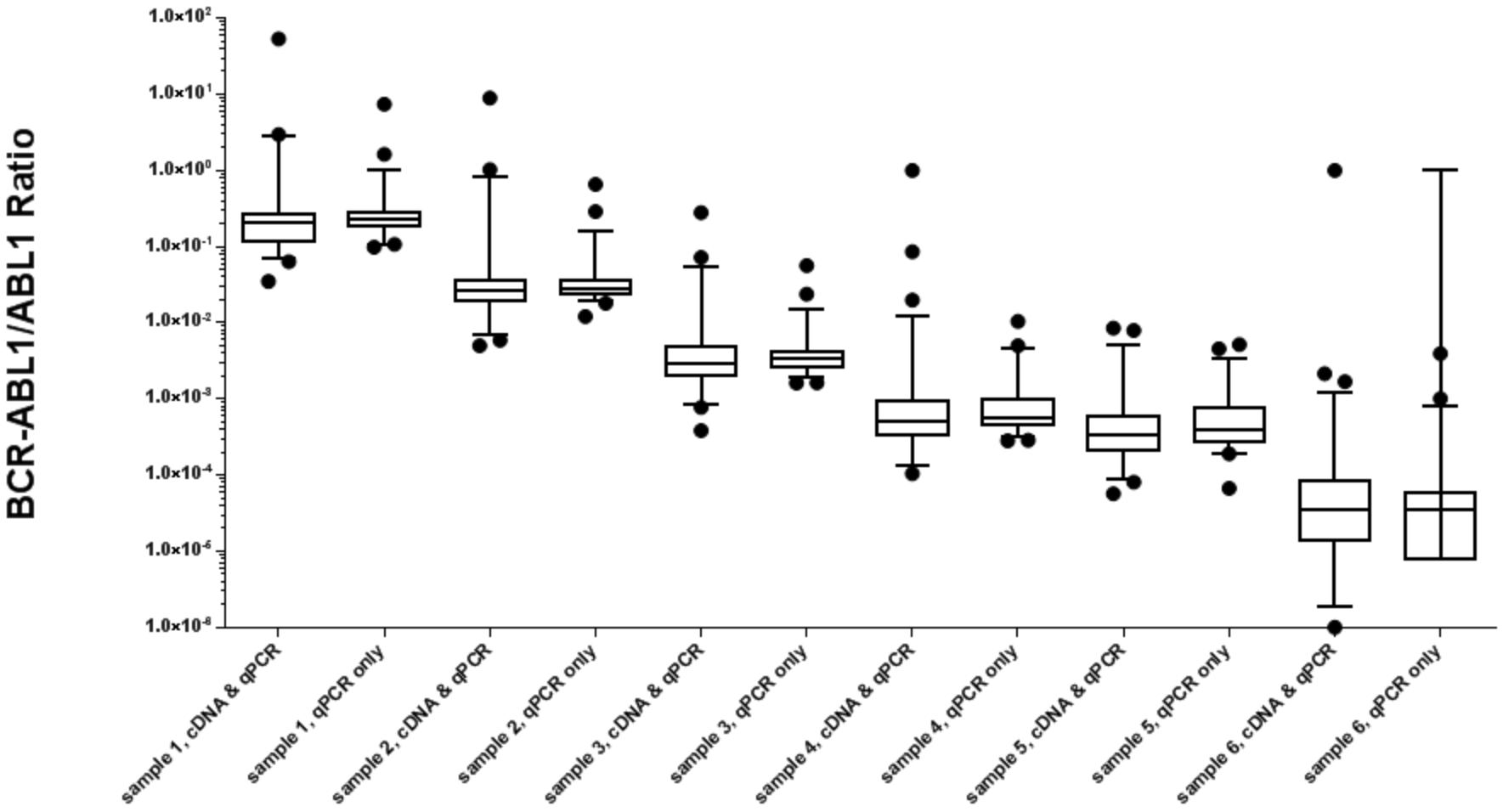
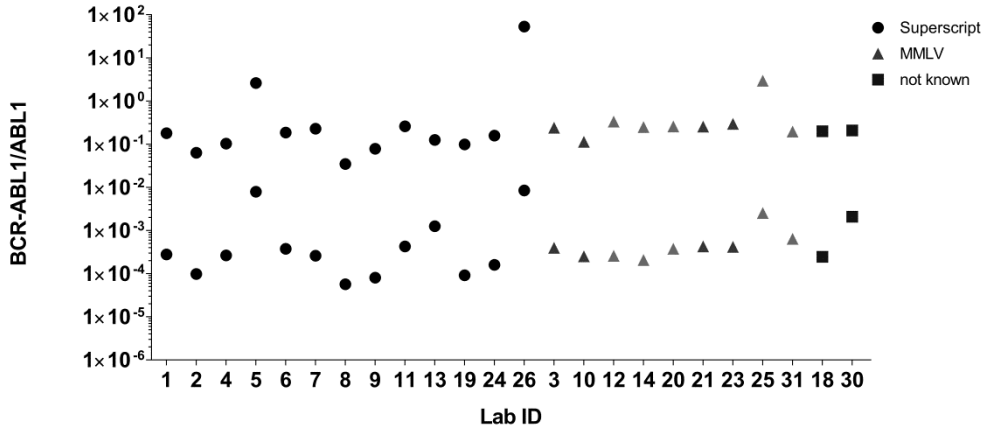


Figure 3: Inter-laboratory variability of *BCR-ABL1/ABL1* quantitation in relation to method of cDNA synthesis and cycler platform

A: local cDNA synthesis and BCR-ABL1/ABL1 ratio, enzymes



B: local cDNA synthesis and BCR-ABL1/ABL1 ratio by cycler type

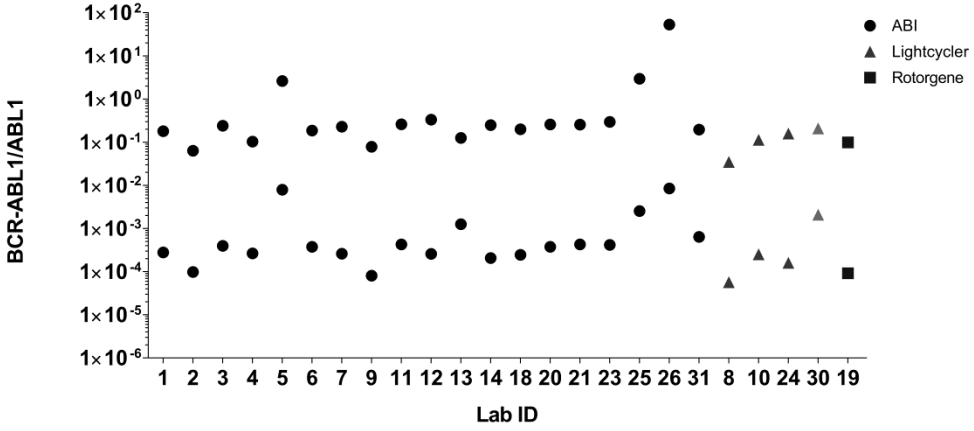
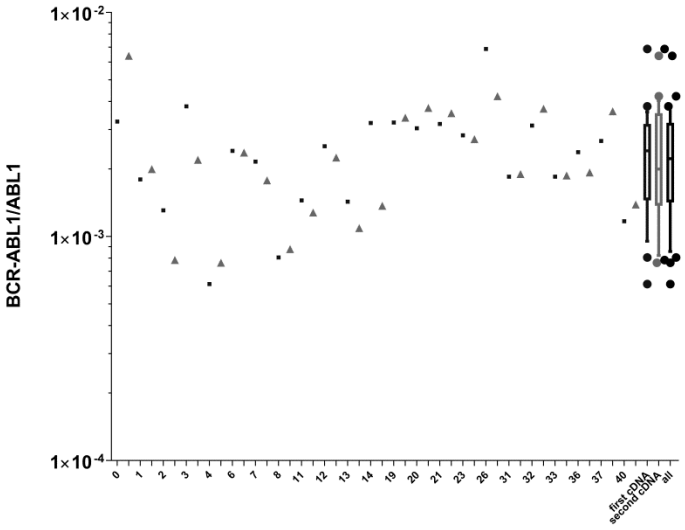


Figure 4: Inter-laboratory variability of e1a2 *BCR-ABL1* quantitation in QC 10 using identical plasmid standards

A: BCR-ABL1/ABL1



B: BCR-ABL1/GUS

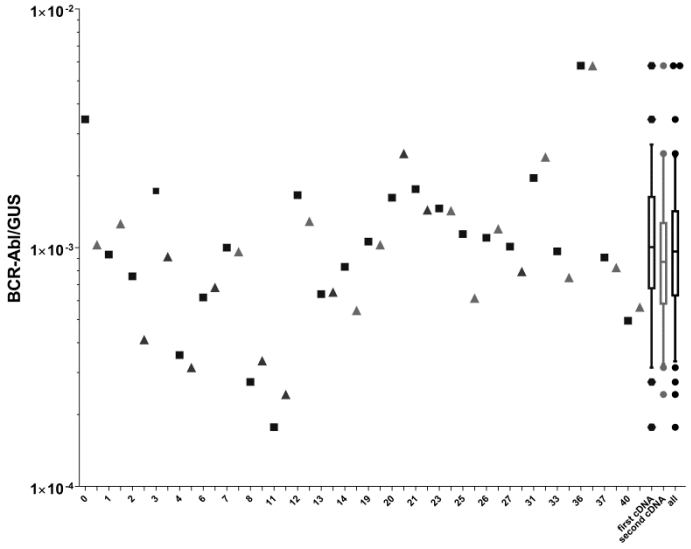
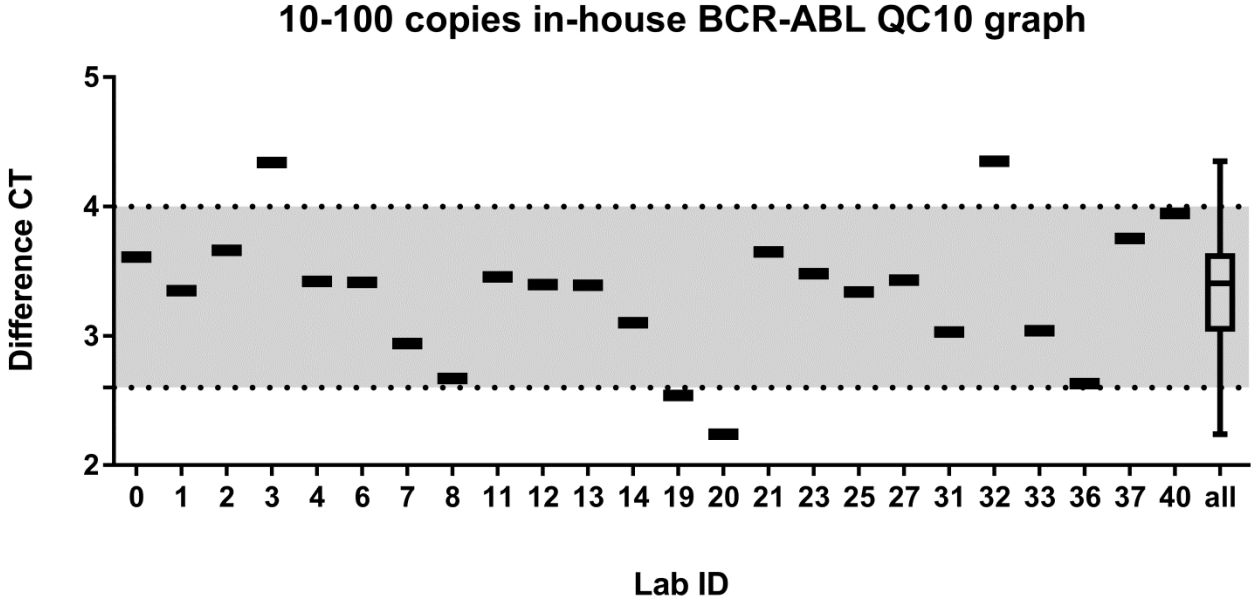


Figure 5: Ability of laboratories to detect *BCR-ABL1* copies within the QR

An example of 154 datasets showing the ability of participating labs to detect 10-100 p190 BCR-ABL copies within the quantitative range defined in QC6 (with in-house plasmids)



4 Labs are outside the quantitative range (2.6 - 4.0)

Table 1: Evaluation of the frequency of false negative and false-positive results

quality control round		QC 1				QC 2			QC 3			QC 4		QC 5		quality control round		QC 6			QC 7	QC 8		QC 9			QC 10		
no. of labs		22*				20			21			21		26		no. of labs		26			-	29		26			25		
no. of returned results		22				20			21			21		25		no. of returned results		26			-	29		25			25		
sample ID		Median	Batch			Median	Batch		Median	Batch		Median	Batch	Median	Batch	sample ID		Median	Batch			Median	Batch	Median	Batch		Median	Batch	
			A	B	C		A	B		A	B		A		A				A	B	-		A		A	B		A	B
1	pos	1,93E-01	22	22	22	1,85E-01	20	20	2,16E-01	18	18	2,54E-01	18	2,12E-01	21	1	pos	2,34E-01	26	25	-	-	-	-	-	-	1,33E-01	23	23
	false neg		0	0	0		0	0		0	0		0		0		false neg		0	0	-		-		-	-		0	0
	n.a.		0	0	0		0	0		3	3		3		4		n.a.		0	1	-		-		-	-		2	2
2	pos	4,41E-02	22	22	22	3,87E-02	20	20	1,10E-01	18	18	-	-	-	-	2	pos	2,84E-02	26	25	-	-	-	-	-	-	-	-	-
	false neg		0	0	0		0	0		0	0		-		-		false neg		0	0	-		-		-	-		-	-
	n.a.		0	0	0		0	0		3	3		-		-		n.a.		0	1	-		-		-	-		-	-
3	pos	2,51E-02	22	22	22	2,21E-02	20	20	2,72E-02	18	18	3,31E-02	18	2,53E-02	21	3	pos	3,39E-03	26	25	-	3,41E-02	28	1,70E-02	20	21	1,73E-03	23	23
	false neg		0	0	0		0	0		0	0		0		0		false neg		0	0	-		0		0	0		0	0
	n.a.		0	0	0		0	0		3	3		3		4		n.a.		0	1	-		1		5	4		2	2
4	pos	2,57E-03	21	22	22	2,40E-03	20	20	7,03E-04	18	18	4,25E-03	18	3,31E-03	21	4	pos	5,71E-04	26	25	-	4,12E-04	28	-	-	-	1,39E-03	22	21
	false neg		1	0	0		0	0		0	0		0		0		false neg		0	0	-		0		-	-		1	2
	n.a.		0	0	0		0	0		3	3		3		4		n.a.		0	1	-		1		-	-		2	2
5	pos	2,27E-04	22	22	22	2,56E-04	20	20	1,31E-04	18	17	3,69E-04	18	2,31E-04	18	5	pos	4,03E-04	26	25	-	-	-	3,08E-04	19	21	-	-	-
	false neg		0	0	0		0	0		0	0		0		3		false neg		0	0	-		-		1	0		-	-
	n.a.		0	0	0		0	0		3	4		3		4		n.a.		0	1	-		-		5	4		-	-
6	pos	2,26E-05	22	20	21	2,85E-05	16	17	4,25E-05	17	18	3,09E-05	15	1,72E-05	16	6	pos	3,41E-05	22	23	-	-	-	-	-	-	-	-	-
	false neg		0	2	1		4	3		1	0		3		5		false neg		4	2	-		-		-	-		-	-
	n.a.		0	0	0		0	0		3	3		3		4		n.a.		0	1	-		-		-	-		-	-
7	false pos	0,00E+00	1	1	4	0,00E+00	0	0	0,00E+00	2	2	0,00E+00	0	0,00E+00	1	7	false pos	0,00E+00	3	3	-	0,00E+00	4	0,00E+00	3	0	0,00E+00	0	0
	neg		21	21	18		20	20		15	13		18		20		neg		23	22	-		23		17	21		23	23
	n.a.		0	0	0		0	0		3	3		3		4		n.a.		0	1	-		1		5	4		2	2

***Lab 0 : one analyses more, a total of 67 analyses is available**

Table 2: Guidelines for Assessment of p190$^{BCR-ABL1}$: Data requirements and definitions

Parameter	Criteria
Standard curve	• ≥ 3 standard points for ABL1/GUS, at least in duplicate • ≥ 5 standard points for BCR-ABL1, at least in duplicate • covering 10 – 10^5 copies for BCR-ABL1 • slope: between -3.2 and -3.7 • correlation coefficient $\underline{>}$ 0.98
Quantitative range	Defined by the lowest dilution of standards with: • a reproducible Delta-CT of all replicates ≤ 1,5 CT • a specific amplification (determined by shape and multicomponent graph) • mean CT value within 2.6 – 4.0 CT for 1 log difference and 0.5 – 1.5 CT for two-fold dilutions
Sensitivity of the assay	Lowest dilution of standards positive with • a specific amplification curve • at least one positive replicate (CT with intercept +1) independent of the reproducibility and the Delta-CT difference
MRD-Positivity	• at least one replicate is positive within Intercept +1 • with a specific amplification curve • any amplification beyond this point: undetermined, not clearly negative, questionable positive
MRD quantitation	• if the Delta-CT of the replicates is ≤ 1.5 • if the mean CT value is less than or equal to the highest CT value of the quantitative range • MRD-positive samples outside the quantitative range: scored as positive, outside quantitative range.
Optimal sample quality	• ≥10.000 ABL1 copies • The minimum ABL1 amount for calculating a ratio should be evaluated by further patient data.