Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia

OPEN

N C P Cross, H White, D Colomer, H Ehrencrona, L Foroni, E Gottardi, T Lange, T Lion, K M Polakova, S Dulucq, G Martinelli, E O Leibundgut, N Pallisgaard, G Barbany, T Sacha, R Talmaci, B Izzo, G Saglio, F Pane, M C Müller, A Hochhaus

Cite this article as: N C P Cross, H White, D Colomer, H Ehrencrona, L Foroni, E Gottardi, T Lange, T Lion, K M Polakova, S Dulucq, G Martinelli, E O Leibundgut, N Pallisgaard, G Barbany, T Sacha, R Talmaci, B Izzo, G Saglio, F Pane, M C Müller, A Hochhaus, Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia, *Leukemia* accepted article preview 5 February 2015; doi: 10.1038/leu.2015.29.

Received 16 January 2015; accepted 21 January 2015; Accepted article preview online 5 February 2015

Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia

Nicholas C.P. Cross[1,2], Helen White[1,2], Dolors Colomer[3], Hans Ehrencrona[4], Letizia Foroni[5], Enrico Gottardi[6], Thoralf Lange[7], Thomas Lion[8], Katerina Machova Polakova[9], Stéphanie Dulucq [10], Giovanni Martinelli[11], Elisabeth Oppliger Leibundgut[12], Niels Pallisgaard[13], Gisela Barbany[14], Tomasz Sacha[15], Rodica Talmaci[16], Barbara Izzo[17], Giuseppe Saglio[6], Fabrizio Pane[17,18], Martin C. Müller[19], Andreas Hochhaus [20]

[1]Faculty of Medicine, University of Southampton, Southampton, UK
[2]National Genetics Reference Laboratory (Wessex), Salisbury, UK
[3]Hematopathology Unit, Hospital Clinic, IDIBAPS, Barcelona, Spain
[4]Department of Clinical Genetics, Laboratory Medicine, Lund University, Lund, Sweden
[5]Imperial Molecular Pathology Laboratory, Centre for Haematology, Imperial College London, London, UK
[6]Department of Clinical and Biological Science, University of Turin, Turin, Italy
[7]Abteilung für Hämatologie und internistische Onkologie, Universität Leipzig, Leipzig, Germany
[8]Children´s Cancer Research Institute/LabDia Labordiagnostik and Medical University, Vienna, Austria
[9]Institute of Hematology and Blood Transfusion, Prague, Czech Republic
[10]Laboratoire Hematologie, CHU Bordeaux, Hematopoiese Leucemique et Cibles Therapeutiques, INSERM U1035, Université Bordeaux, France
[11]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Italy
[12]Molecular Diagnostics Laboratory, Department of Hematology, University Hospital Bern, Bern, Switzerland
[13] Klinisk Biokemi, Vejle Sygehus, Vejle, Denmark
[14] Department of Molecular Medicine and Surgery, Clinical Genetics, Karolinska Institutet, Stockholm, Sweden

[15]Hematology Department, Jagiellonian University, Krakow, Poland
[16]Hematology Department - Fundeni Clinical Institute, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania
[17]Department of Clinical Medicine and Surgery, University "Federico II" of Naples, Naples, Italy
[18]CEINGE – Biotecnologie Avanzate, Naples, Italy
[19]III. Medizinische Klinik, Universitätsmedizin Mannheim, Mannheim, Germany
[20]Abteilung Hämatologie/Onkologie, Klinik für Innere Medizin II, Universitätsklinikum Jena, Jena, Germany

Correspondence to:

Professor Nicholas C.P. Cross
Wessex Regional Genetics Laboratory
Salisbury District Hospital
Salisbury SP2 8BJ, UK

Tel +44 1722 429080
Fax +44 1722 331531
Email: ncpc@soton.ac.uk

Abstract

Treatment of chronic myeloid leukemia (CML) with tyrosine kinase inhibitors has advanced to a stage where many patients achieve very low or undetectable levels of disease. Remarkably, some of these patients remain in sustained remission when treatment is withdrawn, suggesting that they may be at least operationally cured of their disease. Accurate definition of deep molecular responses (MR) is therefore increasingly important for optimal patient management and comparison of independent data sets. We previously published proposals for broad standardized definitions of MR at different levels of sensitivity. Here we present detailed laboratory recommendations, developed as part of the European Treatment and Outcome Study (EUTOS) for CML, to enable testing laboratories to score MR in a reproducible manner for CML patients expressing the most common *BCR-ABL1* variants.

Introduction

Molecular monitoring provides important prognostic information for individual chronic myeloid leukemia (CML) patients undergoing therapy, and international treatment recommendations incorporate specific time-dependent molecular milestones to help determine whether a patient is responding optimally or not.[1, 2] Molecular measurements are made by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) to estimate the amount of *BCR-ABL1* mRNA relative to an internal reference gene, most commonly *ABL1*, *GUSB,* or *BCR*.[3, 4] The results are expressed on an International Scale (IS) as a percentage, with 100% BCR-ABL^{IS} corresponding to the International Randomized Study of Interferon and STI571 (IRIS) study standardized baseline

and 0.1% BCR-ABL^{IS} being defined as a major molecular response (MMR or MR^3; 3 log reduction from the standardized baseline).[3] Expression of results on the IS depends on each testing laboratory either having obtained a laboratory-specific conversion factor (CF) by sample exchange with an established reference laboratory, or by using kits and reagents that have been calibrated to the World Health Organisation International Genetic Reference Panel for quantitation of *BCR-ABL1* mRNA.[4-9]

Efforts to standardize molecular monitoring to the IS focused initially on detectable residual disease and in particular whether a patient had or had not achieved particular milestones, for example 10% BCR-ABL^{IS} or 0.1% BCR-ABL^{IS} at various time points. However with longer follow up, it became apparent that many patients treated with imatinib achieved deeper levels of response, with *BCR-ABL1* becoming undetectable in a minority of cases.[10] This, along with the fact that second generation TKIs produce faster and deeper responses compared to imatinib,[11, 12] prompted the need for robust, standardized definitions of deep molecular response (MR). Such definitions are particularly important in the context of studies that are enrolling patients with sustained deep responses into treatment-free protocols .[13, 14]

We previously published proposals for broad standardized definitions of MR at different levels of sensitivity (MR^4, MR$^{4.5}$ etc; collectively referred to as 'deep MR') which were endorsed by the European LeukemiaNet in their most recent recommendations for treatment of CML patients.[1, 15] These broad definitions, however, and clinical studies that have been published to date, do

not provide the technical details and interpretation to enable laboratories to categorize patients in a standardized fashion. As part of the European Treatment and Outcome Study (EUTOS) we have developed laboratory proposals, as detailed below, to enable testing laboratories to define MR in a reproducible manner. These proposals were developed by consensus over several meetings and are described in detail in this paper, along with several examples. The terminology employed is based on the recommendations of the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines[16] and the proposal focuses on qPCR assays for the most common *BCR-ABL1* variants (e13a2 and/or e14a2; 97% of CML patients) that use an external plasmid calibrator to estimate numbers of target molecules.

Reference genes other than *ABL1*

The published definitions of MR focus on the use of *ABL1* as a reference gene since this is used by the majority of laboratories worldwide.[15] Of the principal alternative reference genes,[3] *GUSB* is used by a significant minority of European laboratories whereas *BCR* is used primarily in Australasia and some US laboratories. We have focused here on extending the MR definitions when *BCR-ABL1* is undetectable to include *GUSB*; further work will be required to extend these definitions to include *BCR*.

To determine the correspondence between *ABL1* and *GUSB*, we collected data from three centres that routinely analysed the expression of both genes in parallel. We focused on CML

samples that were <10% BCR-ABL^{IS} and had >10,000 *ABL1* copies. Of 1567 samples, the median ratio of *GUSB*/*ABL1* was 2.4 in the same volume of cDNA and therefore we consider that, for the purpose of defining deep MR, 10,000 *ABL1* transcripts are equivalent to 24,000 *GUSB* transcripts. The previously published[15] definitions of MR can therefore be expanded as follows:

- MR^4 (≥4-log reduction from IRIS baseline) = either (i) detectable disease ≤0.01% BCR-ABL^{IS} or (ii) undetectable disease in cDNA with 10,000-31,999 *ABL1* transcripts or 24,000-76,999 *GUSB* transcripts
- MR$^{4.5}$ (≥4.5-log reduction from IRIS baseline) = either (i) detectable disease ≤0.0032% BCR-ABL^{IS} or (ii) undetectable disease in cDNA with 32,000-99,999 *ABL1* transcripts or 77,000-239,999 *GUSB* transcripts
- MR^5 (≥5-log reduction from IRIS baseline) = either (i) detectable disease ≤0.001% BCR-ABL^{IS} or (ii) undetectable disease in cDNA with ≥100,000 *ABL1* transcripts ≥240,000 *GUSB* transcripts

Although *GUSB* laboratories may use these definitions, we suggest that they should ideally derive their own correspondence between *ABL1* and *GUSB* (or other reference gene) using at least 50-100 remission (<10% BCR-ABL^{IS}) samples to derive their own cut offs for different MR levels. Prior to making this comparison, the amplification conditions should be optimized and in particular the amplification efficiency for the two genes should be the same. This can be achieved easily for *ABL1*, *GUSB* and *BCR* (and *BCR-ABL1*) using the ERM-AD623 plasmid.[17] For

laboratory developed tests we further recommend that ERM-AD623 is used directly as a qPCR calibrator for routine analysis or indirectly as a calibrator for in house plasmid dilutions.

Defining detectable and undetectable disease

There are several ways in which testing laboratories differ in how they define disease as detectable or undetectable. For individual amplification reactions and runs, we recommend that the established Europe Against Cancer (EAC) criteria are used[18]. In particular:

- The cut-off for positivity should correspond to a quantification cycle (Cq) of intercept + 1 (which should generally lead to cut-offs of 41 – 42 Cq). In other words, samples with a Cq higher than intercept + 1 should be considered as undetectable.
- The 'no template control' wells and reagent blanks should ideally not cross the threshold at any point but should certainly be at least 2 Cq above the intercept Cq for that run. If this is not the case then the run must be considered as failed.

A major variable between centres is the number of replicate assays that are performed for each sample and the way that those replicates are considered to yield the final result. Typically, both *BCR-ABL1* and the reference gene are tested in duplicate, although some centres perform triplicate assays and some only perform single assays. If replicate assays are performed for *BCR-ABL1* (as recommended from RNA[19, 20] or cDNA[21] to help improve the accuracy of results) and any of the individual replicates are positive according to the criteria above, we recommend that

the final result is considered as positive, i.e. detectable disease. Even when testing in triplicate and two replicates are scored as undetectable and one is scored as detectable, the overall result should be scored as detectable or positive.

The EAC defines assay sensitivity by using normalized copy number and ΔΔCt methods, both of which relate the level of MRD to pretreatment levels for individual patients.[22] This is not compatible with the IS in CML, which relates MRD levels to the IRIS standardized baseline, and therefore an alternative approach is required.

Scoring MR when disease is detectable

In general, measurable residual disease[23] should be assigned a value on the IS and scored as MR^4 if ≤0.01% $BCR\text{-}ABL^{IS}$, $MR^{4.5}$ if ≤0.0032% $BCR\text{-}ABL^{IS}$ etc, provided that the sample fulfils minimum quality criteria, i.e *ABL1* ≥10,000 or *GUSB* ≥24,000 in each replicate.[21] If replicate analyses are performed and the values between replicates are comparable[21] then the number of *BCR-ABL1* and reference gene transcripts should be the total value across replicates and the final result expressed on the IS, i.e. [(sum of *BCR-ABL1* copies)/(sum of reference gene copies)] x conversion factor x 100 (see examples 1-3). Since the reference gene in this context is used to estimate the amount of cDNA tested for *BCR-ABL1*, any difference in the number of replicates performed for *BCR-ABL1* and the reference gene will need to be taken into account, (see example 4). In addition, we recommend that for scoring $MR^{4.5}$, the total reference gene number should be 32,000-99,999 *ABL1* transcripts or 77,000-239,999 *GUSB* transcripts regardless of

whether the disease is detectable or undetectable. For scoring MR[5] the total reference gene number should be *ABL1* ≥100,000 or *GUSB* ≥240,000 (Table 1; see example 5).

Many centers score positive samples with a Cq higher than that of the lowest plasmid standard as 'low level positive', positive outside the quantifiable range (POQR), '<10 *BCR-ABL1*' (if the lowest standard is 10), '*<4 BCR-ABL1*' (if the lowest standard is 4, etc). Indeed some guidelines specifically recommend that values should not be estimated if they require extrapolation beyond the span of the standard plasmid calibration curve.[21] This presents a problem for scoring low levels of disease and means, for example, that a laboratory using 10 as the lowest standard and a CF = 1 would need to achieve *ABL1* reference gene values of 100,000 or greater to be able to score a sample with low level detectable disease as MR[4] and a value of ≤320,000 to score a similar sample as MR[4.5] (<10 *BCR-ABL1*/320,000 *ABL1* = ≤0.0032% *BCR-ABL1*). Despite the significant errors in quantifying small numbers of target molecules, we suggest that all level positive replicates should be assigned a specific number of *BCR-ABL1* transcripts by extrapolating below the lowest plasmid standard.

Testing laboratories have generally not rigorously determined their in house limit of detection (LoD; defined as the lowest concentration of target that can be detected with 95% confidence) for *BCR-ABL1* transcripts. One reason for this is that standardized reagents have not been available to perform LoD analysis in a reproducible manner. Now, with the availability of the ERM-AD623 plasmid[17] and other calibration reagents[8], we recommend that laboratories

specifically measure[24] and optimize their *BCR-ABL1* LoD. Clearly, the accuracy and precision by which MR can be scored critically depends on the *BCR-ABL1* LoD being maximized. A laboratory with a poor LoD may score samples as undetectable and deep MR, whereas a laboratory with an optimized LoD may detect *BCR-ABL1* in the same sample and score it as not deep MR. Several studies have indicated that qPCR can be routinely optimized to detect single target molecules.[25-27] Assuming that a single *BCR-ABL1* cDNA target can be detected and that there is no background signal (Limit of Blank = 0), the LoD is given by the Poisson distribution as 3 *BCR-ABL1* targets, since for a sample with an average of 3 targets/unit volume, there is a 95% chance that any unit volume will contain at least one target (Figure 1). Thus, we recommend that any replicate scored as positive should be assigned a value of ≥3 *BCR-ABL1* copies, i.e. positive replicates with estimated copy numbers of <3 should be scored as 3 (see examples 6-8). Alternative technologies, e.g. digital PCR, are likely to be more accurate than qPCR for estimating small numbers of target molecules and may well become the method of choice for more accurate definition of low level positive disease.[28, 29]

Scoring MR when disease is undetectable

Analysis of multiple replicates can increase the sensitivity of detection simply by increasing the amount of sample that is tested. This approach has been used to design very sensitive assays to detect *BCR-ABL1* by qRT-PCR in healthy individuals,[30] for genomic DNA based tests for *BCR-ABL1* in CML[31-33], for detection of minimal residual disease in lymphoid disorders[25] and for other applications such as non-invasive prenatal testing[34]. When *BCR-ABL1* is undetectable in all

replicates from the same sample we recommend that the final result is given as [undetectable *BCR-ABL1*]/[sum of reference gene in all the replicates]. We suggest that for routine analysis a maximum of two or three replicates are performed (examples 9-11) although for specific studies it may be desirable to perform more replicates. Stringent quality criteria are essential, specifically replicates with <10,000 *ABL1* or <24,000 *GUSB* transcripts should be considered as inevaluable for determining deep MR (examples 12 and 13), and laboratories should maximize their LoD for *BCR-ABL1* to avoid false negative results. As above, any difference in the number of replicates performed for *BCR-ABL1* and the reference gene should be taken into account (example 14).

Examples

(i) *BCR-ABL1* detected in at least one replicate

- Example 1 (Lab conversion factor = 0.8):

  - *BCR-ABL1* replicate 1: detectable in 2µl cDNA, estimated 7 copies

  - *BCR-ABL1* replicate 2: detectable in 2µl cDNA, estimated 3 copies

  - *ABL1* replicate 1: 24,000 copies in 2µl cDNA

  - *ABL1* replicate 2: 28,000 copies in 2µl cDNA

Result = (sum *BCR-ABL1* = 10)/(sum *ABL1* = 52,000) x0.8 x100 =

0.015% = MMR but not MR^4

- Example 2 (Lab conversion factor = 1.8):

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: detectable in 5µl cDNA, estimated 3 copies

  - *GUSB* replicate 1: 43,000 copies in 5µl cDNA

  - *GUSB* replicate 2: 49,000 copies in 5µl cDNA

Result = (sum *BCR-ABL1* = 3)/(sum *GUSB* = 92,000) x1.8 x100 =

0.0059% = MR^4

Comment: Testing laboratories use different amounts of RNA to make cDNA, make different volumes of cDNA and use different volumes of cDNA for individual qPCR assays. The number of reference gene transcripts should be estimated in the same volume of cDNA used to test for *BCR-ABL1*. The use of other reference genes, e.g. *BCR*, is possible but the numbers of transcripts required to define different levels of MR remain to be determined.

- Example 3 (Lab conversion factor = 0.5):

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: detectable in 5µl cDNA, estimated 3 copies

  - *ABL1* replicate 1: 9,000 copies in 5µl cDNA

- *ABL1* replicate 2: 8,000 copies in 5µl cDNA

Result = inevaluable for MR

Comment: Although the [(sum of *BCR-ABL1*)/(sum of reference gene)] x conversion factor x 100 is less than 0.01%, the sample should be considered as inevaluable for assessment of MR as the *ABL1* copy number in each replicate is <10,000.

- Example 4 (Lab conversion factor = 0.8):

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA
  - *BCR-ABL1* replicate 2: undetectable in 5µl cDNA
  - *BCR-ABL1* replicate 3: detectable in 5µl cDNA, estimated 4 copies
  - *ABL1* replicate 1: 14,000 copies in 5µl cDNA
  - *ABL1* replicate 2: 15,000 copies in 5µl cDNA

Result = (sum *BCR-ABL1* = 4)/(sum *ABL1* = 29,000 x1.5) x0.8 x100 = 0.0074% = MR^4

Comment: The sum of the reference gene copy number is multiplied by 1.5 (equivalent to multiplying the mean copy number x 3) because only two replicates were performed for the

reference gene whereas three replicates were performed for *BCR-ABL1*. In general we consider that it is better to perform the same number of replicates for *BCR-ABL1* and the reference gene.

- Example 5 (Lab conversion factor = 0.25):

  - *BCR-ABL1* replicate 1: undetectable in 2µl cDNA

  - *BCR-ABL1* replicate 2: detectable in 2µl cDNA, estimated 3 copies

  - *ABL1* replicate 1: 12,000 copies in 2µl cDNA

  - *ABL1* replicate 2: 14,000 copies in 2µl cDNA

Result = (sum *BCR-ABL1* = 3)/(sum *ABL1* = 26,000) x0.25 x100 = 0.0029%; sum of *ABL1* <32,000 = MR^4

Comment: Although the [(sum of *BCR-ABL1*)/(sum of reference gene)] x conversion factor x 100 is <0.0032%, the total *ABL1* value is less than 32,000 and should thus be considered as MR^4. Considering the extreme examples of 31,999 *ABL1* transcripts and either 0 or 3 *BCR-ABL1* transcripts, it is apparent that this discrepancy only arises if the CF is <0.35 if using *ABL1* as a reference gene (or <0.82 if using *GUSB*).

- Example 6 (Lab conversion factor = 0.8):

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: detectable in 5µl cDNA, estimated 2 copies

- *ABL1* replicate 1: 18,000 copies in 5µl cDNA

- *ABL1* replicate 2: 16,500 copies in 5µl cDNA

Result = (sum *BCR-ABL1* = 3)/(sum *ABL1* = 34,500) x0.8 x100 =

0.007% = MR[4]

Comment: Each positive replicate should be assigned a value of ≥3 copies and therefore the second *BCR-ABL1* replicate is scored as 3 copies.

- Example 7 (Lab conversion factor = 0.8):

  - *BCR-ABL1* replicate 1: detectable in 5µl cDNA, estimated 2 copies

  - *BCR-ABL1* replicate 2: detectable in 5µl cDNA, estimated 1 copy

  - *ABL1* replicate 1: 18,000 copies in 5µl cDNA

  - *ABL1* replicate 2: 16,500 copies in 5µl cDNA

Result = (sum *BCR-ABL1* = 6)/(sum *ABL1* = 34,500) x0.8 x100 =

0.014% = MMR but not MR[4]

Comment: Each positive replicate should be assigned a value of ≥3 copies and therefore each *BCR-ABL1* replicate is scored as 3 copies.

- Example 8 (Lab conversion factor = 0.8):

  - *BCR-ABL1* replicate 1: detectable in 5µl cDNA, estimated 2 copies

  - *BCR-ABL1* replicate 2: detectable in 5µl cDNA, estimated 5 copies

  - *BCR-ABL1* replicate 3: detectable in 5µl cDNA, estimated 7 copies

  - *ABL1* replicate 1: 34,000 copies in 5µl cDNA

  - *ABL1* replicate 2: 38,500 copies in 5µl cDNA

  - *ABL1* replicate 3: 32,500 copies in 5µl cDNA

  Result = (sum *BCR-ABL1* = 15)/(sum *ABL1* = 105,000) x0.8 x100 = 0.011% = MMR but not MR[4]

Comment: Each positive replicate should be assigned a value of ≥3 copies.

(ii) *BCR-ABL1* undetected in all replicates

- Example 9:

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: undetectable in 5µl cDNA

  - *ABL1* replicate 1: 16,500 copies in 5µl cDNA

  - *ABL1* replicate 2: 18,000 copies in 5µl cDNA

Result = undetectable *BCR-ABL1* in 34,500 *ABL1* = MR$^{4.5}$

- Example 10:

  - *BCR-ABL1* single analysis: undetectable in 5µl cDNA

  - *ABL1* single analysis: 45,000 copies in 5µl cDNA

Result = undetectable *BCR-ABL1* in 45,000 *ABL1* = MR$^{4.5}$

Comment: although single analyses are performed by some centres, replicate assays from RNA or cDNA improves the accuracy of results[19-21].

- Example 11:

  - *BCR-ABL1* replicate 1: undetectable in 2µl cDNA

  - *BCR-ABL1* replicate 2: undetectable in 2µl cDNA

  - *BCR-ABL1* replicate 3: undetectable in 2µl cDNA

  - *ABL1* replicate 1: 24,000 copies in 2µl cDNA

  - *ABL1* replicate 2: 22,500 copies in 2µl cDNA

  - *ABL1* replicate 3: 24,000 copies in 2µl cDNA

Result = undetectable *BCR-ABL1* in 70,500 *ABL1* = MR$^{4.5}$

- Example 12:

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: undetectable in 5µl cDNA

  - *ABL1* replicate 1: 7,000 copies in 5µl cDNA

  - *ABL1* replicate 2: 8,000 copies in 5µl cDNA

Result = inevaluable for MR as *ABL1* <10,000 in each replicate

- Example 13:

  - *BCR-ABL1* replicate 1: undetectable in 5µl cDNA

  - *BCR-ABL1* replicate 2: undetectable in 5µl cDNA

  - *ABL1* replicate 1: 6,000 copies in 5µl cDNA

  - *ABL1* replicate 2: 14,000 copies in 5µl cDNA

Result = inevaluable for MR

Comment: One replicate is <10,000 *ABL1* and hence the sample should be considered as inevaluable for MR. Since the two *ABL1* replicates are discordant the reference gene qPCR could be repeated.

- Example 14:

- *BCR-ABL1* replicate 1: undetectable in 2µl cDNA

- *BCR-ABL1* replicate 2: undetectable in 2µl cDNA

- *BCR-ABL1* replicate 3: undetectable in 2µl cDNA

- *ABL1* replicate 1: 16,500 copies in 2µl cDNA

- *ABL1* replicate 2: 18,000 copies in 2µl cDNA

Result = undetectable *BCR-ABL1* in (34,500 x1.5 = 51,750 *ABL1*) = MR$^{4.5}$

Comment: The sum of the reference gene copy number is multiplied by 1.5 because only two replicates were performed for the reference gene whereas three replicates were performed for *BCR-ABL1*.

Concluding remarks

The remarkable progress in the treatment of CML has demanded definitions of deep MR that are stretching the technology of molecular monitoring to its limit. The recommendations described here are an attempt to develop standardized laboratory approaches that strike a reasonable balance between scientific accuracy and clinical reality. It should be recognized that there is considerable inherent uncertainty in defining very low levels of disease and that it will be important to continue to look at trends over time in order to recognize sustained MR. Furthermore, the reproducible application of the recommendations depends critically on the ability of testing laboratories to measure absolute numbers of reference gene transcripts in a

comparable manner, as well as their ability to maximize the LoD for *BCR-ABL1* and minimize inter-assay variability. It is obvious that future methodological improvements that increase the amount of sample tested (as determined by the number of reference gene transcripts) will increase the precision and accuracy of scoring MR^4 or $MR^{4.5}$, as well as enabling even deeper levels of MR to be determined.

We recognize that these recommendations may need to be adapted to local requirements and changing technologies. We also recognize that laboratory recommendations in isolation are meaningless and that the critical question is the clinical significance of achieving deep levels of MR. We anticipate that the standardized definitions described here will help to progress clinical studies that aim ultimately to cure CML as well as providing a common framework for reporting routine results.

CONFLICT OF INTEREST

None of the authors have a relevant conflict of interest. This work was supported by the European LeukemiaNet via the 'European Treatment and Outcome Study' (EUTOS) for CML. EUTOS is funded by Novartis but the recommendations described in this paper were developed, finalised and written entirely by the authors.

ACKNOWLEDGEMENT

We would like to thank all the EUTOS laboratory members and other individuals who have contributed to the discussions that have led to the development of these recommendations. GM was supported by NGS-PTL FP7 European Project, and ER regional project.

FIGURE LEGEND

LoD of *BCR-ABL1* detection. The graph shows the Poisson distribution with a mean of 3 *BCR-ABL1* targets per well. The percentage of wells with 0 – 10 targets per well is indicated (20,000 computer generated random datapoints; Minitab version 16, Coventry, UK) and shows that 95% of wells has at least one *BCR-ABL1* target. Since the LoD is defined as the lowest concentration of target that can be detected with 95% confidence, the maximal theoretical LOD is 3 copies.

References

1. Baccarani M, Deininger MW, Rosti G, Hochhaus A, Soverini S, Apperley JF*, et al.* European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood* 2013; **122:** 872-884.

2. O'Brien S, Radich JP, Abboud CN, Akhtari M, Altman JK, Berman E*, et al.* Chronic myelogenous leukemia, version 1.2015. *J Natl Compr Canc Netw* 2014; **12:** 1590-1610.

3. 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:** 28-37.

4. Cross NC. Standardisation of molecular monitoring for chronic myeloid leukaemia. *Best Pract Res Clin Haematol* 2009; **22:** 355-365.

5. 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:** 1925-1930.

6. 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:** 3330-3338.

7. White HE, Matejtschuk P, Rigsby P, Gabert J, Lin F, Lynn Wang Y*, et al.* Establishment of the first World Health Organization International Genetic Reference Panel for quantitation of BCR-ABL mRNA. *Blood* 2010; **116:** e111-117.

8. White HE, Hedges J, Bendit I, Branford S, Colomer D, Hochhaus A*, et al.* Establishment and validation of analytical reference panels for the standardization of quantitative BCR-ABL1 measurements on the international scale. *Clin Chem* 2013; **59:** 938-948.

9. Maute C, Nibourel O, Rea D, Coiteux V, Grardel N, Preudhomme C*, et al.* Calibration of BCR-ABL1 mRNA quantification methods using genetic reference materials is a valid strategy to report results on the international scale. *Clin Biochem* 2014; **47:** 1333-1336.

10. Hehlmann R, Grimwade D, Simonsson B, Apperley J, Baccarani M, Barbui T*, et al.* The European LeukemiaNet: achievements and perspectives. *Haematologica* 2011; **96:** 156-162.

11. Saglio G, Kim DW, Issaragrisil S, le Coutre P, Etienne G, Lobo C*, et al.* Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med* 2010; **362:** 2251-2259.

12. Kantarjian H, Shah NP, Hochhaus A, Cortes J, Shah S, Ayala M*, et al.* Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2010; **362:** 2260-2270.

13. Mahon FX, Rea D, Guilhot J, Guilhot F, Huguet F, Nicolini F*, et al.* Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. *Lancet Oncol* 2010; **11:** 1029-1035.

14. Rousselot P, Charbonnier A, Cony-Makhoul P, Agape P, Nicolini FE, Varet B*, et al.* Loss of major molecular response as a trigger for restarting tyrosine kinase inhibitor therapy in patients with chronic-phase chronic myelogenous leukemia who have stopped imatinib after durable undetectable disease. *J Clin Oncol* 2014; **32:** 424-430.

15. Cross NC, White HE, Muller MC, Saglio G, Hochhaus A. Standardized definitions of molecular response in chronic myeloid leukemia. *Leukemia* 2012; **26:** 2172-2175.

16. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M*, et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 2009; **55:** 611-622.

17. White H, Deprez L, Corbisier P, Hall V, Lin F, al. e. A certified plasmid reference material for the standardisation of BCR-ABL1 mRNA quantification by eal time quantitative PCR *Leukemia* 2014; **in press**.

18. 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:** 2318-2357.

19. Branford S, Hughes T. Diagnosis and monitoring of chronic myeloid leukemia by qualitative and quantitative RT-PCR. *Methods Mol Med* 2006; **125:** 69-92.

20. Jennings LJ, Smith FA, Halling KC, Persons DL, Kamel-Reid S, Molecular Oncology Resource Committee of the College of American P. Design and analytic validation of

BCR-ABL1 quantitative reverse transcription polymerase chain reaction assay for monitoring minimal residual disease. *Arch Pathol Lab Med* 2012; **136:** 33-40.

21. Foroni L, Wilson G, Gerrard G, Mason J, Grimwade D, White HE*, et al.* Guidelines for the measurement of BCR-ABL1 transcripts in chronic myeloid leukaemia. *Br J Haematol* 2011; **153:** 179-190.

22. 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:** 2474-2486.

23. Goldman JM, Gale RP. What does MRD in leukemia really mean? *Leukemia* 2014; **28:** 1131.

24. CLSI. Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline - 2nd Edition (EP17-A2). Wayne, PA (USA): Clinical and Laboratory Standards Institute.; 2012.

25. Morley AA, Latham S, Brisco MJ, Sykes PJ, Sutton R, Hughes E*, et al.* Sensitive and specific measurement of minimal residual disease in acute lymphoblastic leukemia. *J Mol Diagn* 2009; **11:** 201-210.

26. Ladetto M, Lobetti-Bodoni C, Mantoan B, Ceccarelli M, Boccomini C, Genuardi E*, et al.* Persistence of minimal residual disease in bone marrow predicts outcome in follicular lymphomas treated with a rituximab-intensive program. *Blood* 2013; **122:** 3759-3766.

27. Morley AA. Digital PCR: A brief history. *Biomolecular Detection and Quantification* 2014; **1:** 1-2.

28. Goh HG, Lin M, Fukushima T, Saglio G, Kim D, Choi SY*, et al.* Sensitive quantitation of minimal residual disease in chronic myeloid leukemia using nanofluidic digital polymerase chain reaction assay. *Leuk Lymphoma* 2011; **52:** 896-904.

29. Jennings LJ, George D, Czech J, Yu M, Joseph L. Detection and quantification of BCR-ABL1 fusion transcripts by droplet digital PCR. *J Mol Diagn* 2014; **16:** 174-179.

30. Bose S, Deininger M, Gora-Tybor J, Goldman JM, Melo JV. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance

and implications for the assessment of minimal residual disease. *Blood* 1998; **92:** 3362-3367.

31. Sobrinho-Simoes M, Wilczek V, Score J, Cross NC, Apperley JF, Melo JV. In search of the original leukemic clone in chronic myeloid leukemia patients in complete molecular remission after stem cell transplantation or imatinib. *Blood* 2010; **116:** 1329-1335.

32. Ross DM, Branford S, Seymour JF, Schwarer AP, Arthur C, Bartley PA*, et al.* Patients with chronic myeloid leukemia who maintain a complete molecular response after stopping imatinib treatment have evidence of persistent leukemia by DNA PCR. *Leukemia* 2010; **24:** 1719-1724.

33. Bartley PA, Ross DM, Latham S, Martin-Harris MH, Budgen B, Wilczek V*, et al.* Sensitive detection and quantification of minimal residual disease in chronic myeloid leukaemia using nested quantitative PCR for BCR-ABL DNA. *Int J Lab Hematol* 2010; **32:** e222-228.

34. Hromadnikova I, Houbova B, Hridelova D, Voslarova S, Kofer J, Komrska V*, et al.* Replicate real-time PCR testing of DNA in maternal plasma increases the sensitivity of non-invasive fetal sex determination. *Prenat Diagn* 2003; **23:** 235-238.

Table: Summary of reference gene numbers required for scoring deep molecular response

	MR[4]	**MR[4.5]**	**MR[5]**
Minimum sum of reference gene transcripts irrespective of whether *BCR-ABL1* is detected or not[a]	10,000 *ABL1* 24,000 *GUSB*	32,000 *ABL1* 77,000 *GUSB*	100,000 *ABL1* 240,000 *GUSB*
BCR-ABL^{IS} level for positive samples[b]	≤0.01%	≤0.0032%	≤0.001%

[a] numbers of reference gene transcripts in same volume of cDNA that is tested for *BCR-ABL1.* The minimum number in any individual replicate should be 10,000 *ABL1* or 24,000 *GUSB*.

[b] provided that the minimum reference gene copy numbers in the row above are fulfilled