Published in final edited form as:

Leuk Res. 2009 September ; 33(9): 1170–1172. doi:10.1016/j.leukres.2009.03.014.

Real-time Consensus on Relapse Risk in Acute Promyelocytic Leukemia

Robert E. Gallagher, M.D.

Department of Medicine and Oncology, Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY 10467

Keywords

APL; minimal residual disease; real-time RT-PCR; molecular relapse

In acute promyelocytic leukemia (APL), there are two main purposes for minimal residual disease (MRD) monitoring by reverse transcription-polymerase chain reaction (RT-PCR) methods: to assess the risk of hematological relapse (HR) in patients in clinical remission after the completion of consolidation therapy and to identify patients at high risk of HR as a basis for initiating salvage therapy on the premise that this will provide a more favorable outcome while persistent or recurrent disease is still sub-clinical. The primary evidence affirmatively supporting both of these purposes was provided by non-quantitative, conventional RT-PCR (C-PCR), principally in association with Italian GIMEMA clinical trials, using all-trans retinoic acid (ATRA) in combination with anthracycline-based chemotherapy (ABCT) [1,2]. Subsequently, an optimized, standardized version of the C-PCR assay was incorporated into Spanish PETHEMA ATRA/ABCT trials, in which molecular relapse (MR) was regarded as uncensored, i.e., considered in common with HR, for trial outcome evaluation and was prospectively used for treatment modification in some cases [3,4]. Confident use of these C-PCR assays by GIMEMA/PETHEMA investigators was based on the utilization of the standardized C-PCR assay in multi-institutional protocol trial sites and, critically, the confirmation of all positive initial results using a second sample in two central reference laboratories. Confident use was also enhanced by the clear definition of terms: molecular remission, the disappearance of a gel electrophoretic band generated from PML-RAR α mRNA present in a pretreatment patient sample in a post-treatment bone marrow (BM) sample using an assay with a sensitivity of $\sim 10^{-4}$ established by appropriate positive dilution controls; molecular persistence, the detection of the PML-RAR α gel band after the termination of consolidation therapy and confirmation of this signal in a second specimen; MR, the reappearance of the confirmed gel signal from documented molecular remission during maintenance therapy or follow-up [3].

More recently, PETHEMA investigators analyzed the effectiveness of real-time quantitative RT-PCR (RQ-PCR) in a subgroup of PETHEMA patients, including a comparison to C-PCR [5]. A crucial element of RQ-PCR measurement is that the PML-RAR α expression level is assessed as a ratio of its real-time value to a concurrently determined real-time value of a

Corresponding Author: Robert E. Gallagher, M.D., Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY 10467, Phone: 718-920-4180, Fax: 718-798-7474, Email: rgallagh@aecom.yu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

constantly-expressed housekeeping gene, which is referred to as the normalized copy number (NCN) or normalized quotient (NQ) [6]. In the presence of detectable PML-RAR α , this ratio provides an inter-sample normalization for variations in RNA amount/quality and effective conversion into the measured PCR-amplified DNA product. In the absence of detectable PML-RAR α , the housekeeping gene real-time value provides an estimate of the detection sensitivity limit for the negative result. In C-PCR assays, measurements of housekeeping gene controls have usually been much less precise. The comparative PETHEMA study used ABL as a housekeeping gene control, and the NCN was determined by multiplying the PML-RAR α /ABL real-time values ratio by 10^4 , according to a previously standardized procedure [5,7]. The RQ-PCR results during the maintenance and follow-up phases of the PETHEMA trials were that all patients (19/19) with NCN >10 relapsed; 6/18 patients (33%) with NCN 1–10 relapsed (12 false positives with respect to relapse); and no patients (0/101) with NCN <1 relapsed. The C-PCR determinations in the corresponding NCN ranges (positive/total) were: NCN >10, 19/19 positive; NCN 1–10, 8/18 positive (3 false positives; 1 false negative); NCN <1, 2/101 positive (2 C-PCR false positives). These results are consistent with two fundamental differences between the applied RQ-PCR and C-PCR assays: 1) the RQ-PCR assay was ~10-fold more sensitive than the C-PCR assay and 2) C-PCR, which measures the end product of PCR amplification, yields more variable results than RQ-PCR, which measures the beginning of PCR amplification [6]. One principal conclusion from these findings is that in the great majority of patients HR becomes inevitable only at relatively high levels of recurrent sub-clinical disease. This is supported by the consideration that the median NCN in these PETHEMA patients at presentation was ~3000 [5], which provides an estimate that the level of disease in the BM of these patients was 3000/>10, i.e., ≤ 300 -fold lower than an APL blast-filled BM or as little as 10-fold lower than what might be detected by standard cytogenetic analysis. From these considerations, it is apparent why a positive C-PCR assay has been so effective as a predictor of HR despite relatively low detection sensitivity, particularly when a requirement has been to perform a confirmatory assay, which has high probability of excluding false positives from the first assay. A second principal conclusion is that the major advantage of RQ-PCR is that it can discriminate patients at moderate risk of HR from those at very low risk.

In this issue of Leukemia Research, Cassinat and colleagues confirm the above-summarized PETHEMA RQ-PCR findings in the context of French-Belgian-Swiss APL Group (FBSAG) clinical trials using ATRA/ABCT therapy [8]. These investigators also found in two successive studies of 223 and 213 patients that the risk of HR stratified into 3 levels of MRD over a single log₁₀ range, separating patients into groups with no (0%, 0%), mixed (38%, 20%) or near total (100%, 83%) incidences of relapse within a few months of testing. A different housekeeping gene PBGD was used for normalization, which, after multiplication by 10^4 , generated the alternative NCN categories of <10, 10–100 and >100. As in the PETHEMA study, the top category must represent a high level of MRD, based on a cited median NCN value of ~10,000 at HR, i.e., $10^4/>10^2$ or an ~100-fold reduction from full-blown relapse values. A difference from the PETHEMA study is that the lower NCN cut-off value (<10) was based only on RQ-PCR assays that showed low-level presence of PML-RAR α in at least 2 of the 3 replicate wells that are customarily run in RQ-PCR assays. Although these findings support the conclusion that measurable NCN values below the lower cut-off indicate a very favorable prognosis, it cannot be concluded that there is 0% chance of relapse, since those initially detected with higher NCN values must have passed through the lower NCN range during the ascent from a previously-documented undetectable level. Similarly, the statement in this paper that the failure to detect MRD in a second sample after an initial positive test “confirms the absence of detectable MRD” likely underestimates the capacity of this technology to detect low-levels of fluctuating or slowly resolving MRD that may reflect host defense-leukemia cell interactions [9].

Another significant difference between the two studies is that it is apparent that a much larger proportion of the MRD monitoring samples from the FBSAG study than from the PETHEMA study was derived from peripheral blood (PB), although the number of PB vs BM samples is only cited for positive samples [8]. For samples with $\text{NCN} \geq 10$, 7 were PB and 17 were BM; only 4 paired BM and PB were assayed with the finding of 2 false negative PB assays with respect to subsequent relapse. In the PETHEMA study, no significant difference between the NCN values for 130 paired BM and PB samples was found [5], in agreement with a previous study. [9] However, in 7 of the PETHEMA patients who relapsed, there was a 24 – 35 day delay in the detection of MR in the PB of 3 patients [5]. Based on these acknowledged fragmentary data, both studies conclude that BM MRD monitoring is preferable.

An important aspect of the confirmatory FBSAG report is that it demonstrates that the prognostic significance of RQ-PCR determinations can be made across technical platforms with differing housekeeping gene controls if each is adequately standardized within the individual reference laboratory(s). The earlier results of the first North American Intergroup APL trial E2491/INT0129, which also used a variation of ABCT/ATRA therapy, are also in accord with this notion. Using GAPDH as a normalizing gene, a NCN of $\geq 10^{-5}$ (no arbitrary multiplier used) during the post-consolidation maintenance/follow-up period was associated with 77% subsequent HR [9]. The median pretreatment NCN was 10^{-2} , indicating that the 1000-fold lower level corresponded to mixed relapse outcome, which appropriately extrapolates to the intermediate NCN ranges of the PETHEMA (1–10) and FBSAG (10–100) studies. Clearly, a universally standardized platform would be preferable. However, the difficulty in achieving this is illustrated by the experience in chronic myeloid leukemia (CML), in which persistent impediments have been encountered in spite of a massive effort through inter-laboratory exchanges of resources to establish international standardization of a BCR-ABL MRD assay [10]. Further, the PETHEMA and FBSAG studies suggest that such precise standardization may be less critical in APL, since only relatively high levels (10^{-2} – 10^{-3} below pretreatment) of MRD are predictive for impending relapse, whereas in CML accurate measurement of changes as little as one-half log measured at $\geq 10^{-3}$ below pretreatment levels may have prognostic/therapeutic significance. [11]

Although not specifically stated, the consensus provided by the PETHEMA and FBSAG studies provide important guidelines for triggering a therapeutic change based on NCN values. The FBSAG report emphasizes the importance of a confirmatory test, although this would have to be done very quickly for patients with a NCN value >100 , since the majority of these patients relapsed in <2 months [8]. For intermediate NCN values, the serial testing data in the FBSAG report suggests that a reasonable trigger would be a repeat test within 1 – 2 months that shows an increase into the high-risk range. This approach appears to have been taken in a British ATRA/ABCT trial monitored by RQ-PCR (although NCN values are not given in the preliminary report), which triggered arsenic trioxide (ATO) salvage therapy and yielded a HR rate of 3% at 3 years follow-up compared 13% in a historical control trial [12]. The latter study illustrates the potential pay-off for applying RQ-PCR molecular relapse criteria. However, performing such MRD monitoring is enormously demanding and costly, and a looming challenge will be to formulate cost-benefit effective monitoring schedules adapted to therapies that can reduce HR rates to $<10\%$ [13] and that, in the case of ATO-supplemented upfront regimens [14], may obscure ATRA/ABCT risk factors that might be used to stratify the intensity of MRD monitoring [5].

Acknowledgments

REG is supported by Grant R01 CA56771 from the National Cancer Institute, Bethesda, MD, USA

References

1. Diverio D, Rossi V, Avvisati G, De Santis S, Pistilli A, Pane F, et al. Early detection of relapse by prospective reverse transcriptase-polymerase chain reaction analysis of the PML/RAR α fusion gene in patient with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter "AIDA" trial. *Blood* 1998;92:784–789. [PubMed: 9680345]
2. Lo Coco F, Diverio D, Avvisati G, Petti MC, Meloni G, Pogliani EM, et al. Therapy of molecular relapse in acute promyelocytic leukemia. *Blood* 1999;94:2225–2229. [PubMed: 10498592]
3. Sanz MA, Lo Coco F, Martin G, Avvisati G, Rayon C, Barbui T, et al. Definition of relapse risk and role of nonanthracycline drugs for consolidation in patients with acute promyelocytic leukemia: a joint study of the PETHEMA and GIMEMA cooperative groups. *Blood* 2000;96:1247–1253. [PubMed: 10942364]
4. Sanz MA, Martin G, Gonzalez M, Leon A, Rayon C, Rivas C, et al. Risk-adapted treatment of acute promyelocytic leukemia with all-*trans*-retinoic acid and anthracycline monotherapy: a multicenter study by the PETHEMA group. *Blood* 2004;103:1237–1243. [PubMed: 14576047]
5. Santamaria C, Chillon MC, Fernandez C, Martin-Jimenez P, Balanzategui A, Sanz RG, et al. Using quantification of the PML-RAR α transcript to stratify the risk of relapse in patients with acute promyelocytic leukemia. *Haematologica* 2007;92:315–322. [PubMed: 17339180]
6. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. *Genome Res* 1996;6:986–994. [PubMed: 8908518]
7. Gabert J, Beillard E, van der Velden V, Bi W, Grimwade D, Pallisgaard N. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction (RQ-PCR) of fusion gene transcripts for residual disease detection in leukemia - A Europe Against Cancer Program. *Leukemia* 2003;17:2318–2357. [PubMed: 14562125]
8. Cassinat B, de Botton S, Kelaïdi C, Ades L, Zassadowski F, Guillemot I, et al. When can real-time quantitative RT-PCR effectively define molecular relapse in acute promyelocytic leukemia patients? (Results of the French Belgian Swiss APL Group). *Leuk Res*. 2009
9. Gallagher RE, Yeap BY, Bi W, Livak KJ, Beaubier N, Rao S, et al. Quantitative real-time RT-PCR analysis of PML-RAR α mRNA in adult acute promyelocytic leukemia: assessment of prognostic significance in adult patients from intergroup protocol 0129. *Blood* 2003;101:2521–2528. [PubMed: 12468436]
10. 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. [PubMed: 18684859]
11. Press RD, Galderisi C, Yang R, Rempfer C, Willis SG, Mauro MJ, et al. A half-log increase in BCR-ABL RNA predicts a higher risk of relapse in patients with chronic myeloid leukemia with an imatinib-induced complete cytogenetic response. *Clin Cancer Res* 2007;13(20):6136–43. [PubMed: 17947479]
12. Grimwade DJ, Javanovic J, Hills R, Nugent E, Patel Y, Flora R, et al. Evaluation of prospective detection of PML-RAR α and RARA-PML fusion transcripts by real-time quantitative PCR (RQ-PCR) to direct pre-emptive therapy with arsenic trioxide (ATO) in acute promyelocytic leukemia (APL) patients treated in the UK MRC AML15 trial. *Blood* 2007;110:166a–167a.
13. Sanz MA, Tallman MS, Lo-Coco F. Tricks of the trade for the appropriate management of newly diagnosed acute promyelocytic leukemia. *Blood* 2005;105:3019–3025. [PubMed: 15604216]
14. Powell BL, Moser B, Stock W, Gallagher RE, Willman CL, Stone RM, et al. Arsenic trioxide consolidation improves event-free and overall survival among patients with newly diagnosed acute promyelocytic leukemia: North American intergroup protocol C9710. *Ann Hematol* 2008;87:S77–S78.