Q. Regarding the subject of minimal residual disease (MRD), how does one reconcile the discrepant percent of residual blasts among morphology, flow cytometric analysis (in-house), FISH, PCR, and, lastly, the flow cytometric analysis result reported by Children's Oncology Group? We have several cases in which the in-house flow, morphology, and FISH were negative, while the COG report was positive for MRD. What is the acceptable percent range of discrepancy among laboratories for flow cytometric analysis?

A. Minimal residual disease can be detected by a variety of techniques and the same approach can be applied to virtually any disease process, provided there is a distinguishing biomarker to differentiate normal from neoplastic cells. All methods, including morphology (~1 percent threshold of detection), FISH (~0.1 percent threshold of detection), nucleic acid amplification (~0.001 percent threshold of detection), and flow cytometry (~0.01–0.001 percent threshold of detection), have varying sensitivities and sources of interference. Additionally, different methods are better suited to particular hematolymphoid malignancies. Thus, it is increasingly recognized that no single technology can be universally recommended; often a combination of testing modalities is required to accurately determine the level of MRD.

The factors contributing to differences in MRD detection among different labs include not only differences in the method(s) chosen, but also methodologic differences in applying a particular technique, preanalytical differences in obtaining sample (as, for example, in testing different aliquots of bone marrow from different aspirate samples), and variability in the experience and competence of laboratories performing the testing. The intrinsic biologic diversity of neoplasms as well as changes that can occur in phenotype and genotype after therapy further complicate all of these variables. Literature comparing methods can be confusing, and in some cases the confusion is traceable to competency-related differences in using the various techniques or subconscious bias in the study design favoring one technique during the comparative study. Moreover, even in studies showing excellent correlation between methods or laboratories, there are always occasional samples that give inexplicable highly divergent results.

The COG acute lymphoblastic leukemia (ALL) flow cytometry MRD studies are performed in two laboratories highly skilled in flow cytometry, one directed by Brent Wood, MD, PhD, at the University of Washington and the other by Michael J. Borowitz, MD, PhD, at Johns Hopkins. These two labs have standardized procedures that have been validated not only analytically but also clinically to ensure similar diagnostic sensitivity and accuracy between the two testing sites. Dr. Wood's summary of the issues follows, in his own words:

Despite all these potential causes for discordance, in most cases discrepancies in enumeration are ±1 log in my experience once denominator effects are taken into account. More specifically:

**Morphology versus flow.** Morphology has limited sensitivity and specificity and is clearly inferior to flow in most cases. Specifically related to ALL, there are a subset of cases in which the blasts acquire a more mature morphologic appearance after therapy that makes them difficult to distinguish from mature lymphocytes and thus morphology can grossly underestimate them (we saw an in-house example of this recently). I’m not sure whether this is well documented in the literature, however. Different denominators are commonly used (white cell events for flow and all nucleated cells for morphology; COG ALL studies use mononuclear cells but I’m not certain about COG acute myeloid leukemia studies [Loken]), and each is subject to its own bias for enumeration due to hemodilution, distributional, and processing artifacts, particularly in bone marrow.

**FISH versus flow.** Concordance is generally good, but enumeration can be problematic, particularly if cultured cells are used as the starting point for FISH. Some abnormalities seen at diagnosis by FISH
may be not present in residual abnormal blasts after therapy, probably due to therapeutic selection, although this is probably relatively infrequent. We conducted a study directly comparing flow and FISH MRD in a large number of samples from patients post-transplant for AML (Fang M, et al. Cancer. 2011;Sep 16. [Epub ahead of print]), and while the concordance is generally good, there are clearly discrepancies in both directions (more FISH– flow+), but the clinical outcomes appear essentially the same comparing the discordant categories to concordant, that is, MRD detected by either technique is equally bad.

**PCR versus flow.** PCR is clearly one log or so more sensitive than flow when either translocation analysis or IgH/TCR patient-specific primers are used. Enumeration with PCR is problematic and it can be difficult to directly compare qPCR with flow numbers. The molecular abnormalities targeted by PCR assays may not always be retained following therapy, and thus false-negative results occur in a subset of patients. Flow offers a more integrated assessment of abnormality and so may be less dependent on specific molecular lesions, but this has not been rigorously shown. I think laboratory competence in performing each of these assays is a major factor in discordance. Both PCR and FISH are somewhat more objective to interpret, and the technical factors are reasonably well understood and somewhat standardized, so that results for either are less likely to be subject to technical or interpretive artifact. On the other hand, flow cytometry is largely unstandardized, with most labs not having appropriate experience with MRD to be truly competent. There is no widely available proficiency testing or other reference method for comparison (aside from those discussed here, each with its own problems), so labs generate results that are largely unsubstantiated. As a result, I think the quality of MRD data from flow labs is highly variable and subject to significant technical and interpretive artifact, a situation that may not be possible to correct before molecular testing gains the upper hand.

Dr. Borowitz of Johns Hopkins agreed with Dr. Wood’s comments and added the following: “Cost should not be an issue,” he said, “at least not in the sense of duplicating cost, as all our studies are grant funded. One could make the case that it is not appropriate for local institutions to charge patients in addition when the reference laboratories are generating the information needed for clinical decisionmaking per protocol.....”

Dr. Borowitz continued: “As to local institutions not detecting MRD when we do at the reference labs, I don’t find it to be that common among what I would consider good labs that understand what MRD testing is about. There are quantitative differences, to be sure, and these likely relate to the things Dr. Wood talked about, as well as the additional factor of shipping. However, I have also seen anecdotal negative cases from labs that perhaps don’t have much experience where the dot plots suggest to me that they would not have had the capability to detect abnormal populations much below one percent or so.”

He added a valuable warning: “Basically, MRD is a different test from leukemia phenotyping analysis and should probably be treated as such. At some point, it should probably be given its own test code, and someone should develop a separate proficiency testing procedure.”

These experts raise valid issues that warrant further discussion, perhaps in the context of the coming debate of who should appropriately regulate laboratory-developed tests, as even leukemia phenotyping performed in nearly every diagnostic flow cytometry laboratory is clearly an LDT. MRD detection is, by anyone’s definition, a high-complexity LDT and challenges the CAP and Food and Drug Administration to understand how best to regulate the practice to ensure quality testing.

In summary, there is no defined consensus for acceptable discrepancies among the methods of MRD detection, but such discrepancies are worthy of note in a comment because an indication of measurement uncertainty might influence a clinical or therapeutic decision. The expectation for some level of discordance given the above-noted technical variables and the evolving implications for patient management speak to the advantages of a consolidated or integrated hematopathology interpretive report, as distinct from several reports coming from the surgical pathology, molecular pathology, flow cytometry, and cytogenetics labs and expecting the treating physician to integrate the information into the correct disease status or subclassification. Those labs that receive reports from clinical trial reference laboratories, such as the COG, might also benefit from this being a form of proficiency
testing for their own laboratories. However, if clinical decisions are based on the outside result, the appropriateness of performing and billing for the internal study might be questionable.

**Suggested Reading**


Bruce H. Davis, MD  
*Treasurer, International Council for Standardization in Haematology  
President, Trillium Diagnostics  
Bangor, Me.*

Member, CAP Hematology/Clinical Microscopy Resource Committee
Q. In reviewing our reference range for MCH, we have noticed an increased incidence of high MCH with normal RBC parameters on certain pre-admission patients. Controls, calibrations, and proficiencies are all well within acceptable limits. Why should we report MCH at all? MCHC flags us for cold agglutinins and spherocytes.

A. The mean corpuscular hemoglobin (MCH) still has a place in clinical diagnosis. MCH is derived from the ratio of total hemoglobin to the erythrocyte count. With the hemoglobin expressed as g/L and the erythrocyte count expressed as $\times 10^{12}$/L, the results are presented in picograms (that is, $10^{-12}$ g). Therefore, MCH provides an absolute gravimetric measurement of the hemoglobin in the average erythrocyte.

MCHC, however, provides the ratio of the hemoglobin to the volume of the average cell. MCHC is useful in many conditions, such as normochromic normocytic anemias, where the amount of hemoglobin produced parallels changes in the MCV. However, when evaluating conditions with a decreased erythrocyte volume, such as thalassemia or iron deficiency, or increased absolute erythrocyte hemoglobin such as hemochromatosis, changes in erythrocyte volumes create problems in interpretation if the MCV or MCHC are used. For example, as erythrocytes age, the MCV will increase and consequently (as shown by the data from David L. Zwick, MD, Fig. 1) the MCHC will decrease, yet the absolute amount of MCH does not change. Such constancy of the MCH is why decreased MCH is used as a tool for evaluating possible thalassemia, iron deficiency, or hemochromatosis.1–3

References


David F. Keren, MD, Professor of Pathology
Associate Director of Clinical Laboratories
University of Michigan Medical School, Ann Arbor

Member, CAP Hematology/Clinical Microscopy Resource Committee