practical or even feasible. This may be particularly true with FA patient cells, which expand poorly in culture and senesce early. On the other hand, iPS cells offer a much superior platform for sophisticated genetic engineering than any somatic cell, allowing for careful selection and characterization of corrected cells, for example, by screening for safe vector integration sites or by gene targeting by homologous recombination.[4,5] Furthermore, reprogramming itself may induce DNA damage, exacerbated in a background of defective DNA damage repair, but cells that harbor a great mutational load seem to be selected against (accounting for the reduced reprogramming efficiency in FA).[6,7] Finally, reprogramming of FA cells also selects for genetically corrected cells that express the therapeutic transgene, as demonstrated in both the present study and the study by Raya et al,[3] so it may be possible to correct and reprogram simultaneously and screen corrected iPS cell clones at the end of the process to exclude genotoxicity imposed by both the genetic modification and reprogramming at once. Which would be the preferred strategy? Further studies to address which strategy can yield iPS cells with adequate efficiency and quality together with anticipated advances in reprogramming and genetic modification methods will eventually inform the next steps toward translation to the clinic.

Conflict-of-interest disclosure: The author declares no competing financial interests. ■

REFERENCES

1. Müller LUW, Milsom MD, Harris CE, et al. Overcoming reprogramming resistance of Fanconi anemia cells. *Blood*. 2012;119(23):5449-5457.
2. Wang W. Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. *Nat Rev Genet*. 2007;8(10):735-748.
3. Raya A, Rodriguez-Piza I, Guenechea G, et al. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature*. 2009;460(7251):53-59.
4. Papapetrou EP, Lee G, Malani N, et al. Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. *Nat Biotechnol*. 2011;29(1):73-78.
5. Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood*. 2011;118(17):4599-4608.
6. Gore A, Li Z, Fung HL, et al. Somatic coding mutations in human induced pluripotent stem cells. *Nature*. 2011;471(7336):63-67.
7. Hussein SM, Batada NN, Vuoristo S, et al. Copy number variation and selection during reprogramming to pluripotency. *Nature*. 2011;471(7336):58-62.

● ● ● CLINICAL TRIALS

Comment on Pabst et al, page 5367

Be quick, but don't hurry

Elihu Estey UNIVERSITY OF WASHINGTON

In this issue of *Blood*, Pabst et al report that granulocyte-colony stimulating factor (G-CSF) "priming" improves event-free and overall survival (EFS and OS) only in those adults less than 60 years old given escalated doses of cytarabine (ara-C) for treatment of newly diagnosed acute myeloid leukemia (AML).[1]

Efforts to improve the frequently unsatisfactory results after treatment of this disease typically entail other cytotoxins in combination with, or as replacements for, standard daunorubicin (or idarubicin) plus ara-C. Another approach emphasizes noncytotoxic drugs to sensitize ("prime") AML blasts to standard therapy. Use of CXCR4 inhibitors to detach marrow blasts from their protective stroma is a recent example,[2] but a much earlier example was G-CSF. Originally given before and/or during standard induction therapy to place more blasts into S-phase of the cell cycle where sensitivity to such therapy is thought greatest, G-CSF priming has had a checkered 20-year history.[3] A particularly noteworthy study (HOVON-SAKK AML-29) whose authors include some of those from the current study randomized 730 adults less than 60 years old with newly diagnosed AML to receive or not receive G-CSF beginning 1 day before, and continuing during, chemotherapy: cycle 1 = idarubicin, + ara-C at 200 mg/m2 daily ×7, cycle 2 = amsacrine, + ara-C at 1g/m2 twice daily ×12.[4] Although G-CSF generally reduced the risk of relapse, an improvement in EFS (hazard ratio [HR] 0.75, $P = .01$) and OS (HR 0.75, $P = .02$) occurred only in the 72% of patients with intermediate risk cytogenetics. Despite these results, G-CSF priming has not found widespread acceptance.

To Pabst et al's great credit a primary purpose of the current, and larger, study (HOVON-42) was to confirm the findings of AML-29, as well as to see if the OS benefit might be more widespread. HOVON-42 was initially conducted within the context of a randomization to either conventional dose ara-C, given as in AML-29, or escalated dose ara-C: cycle 1 = 1g/m2 twice daily ×10, cycle 2 = 2g/m2 twice daily days 1, 2, 4, and 6. Within each of these groups patients were randomized to +/− G-CSF, given during each cycle's chemotherapy. Nine hundred seventeen patients were randomized to +/− G-CSF with 709 receiving conventional dose and 207 escalated dose ara-C. Despite striking similarities between the conventional-dose ara-C arms of AML-29 and HOVON-42, the latter could not reproduce the decrease in relapse risk seen generally in the G-CSF arm of the former, nor the improvement in EFS and OS observed in the intermediate-risk cytogenetic group when given G-CSF (HRs 0.95 and 1.01, respectively, in HOVON-42). There was, however, the above-noted improvement in EFS (HR 0.59, $P = .003$) and OS (HR 0.65, $P = .012$), due primarily to less risk of relapse, in the escalated dose ara-C group given G-CSF.

Pabst and colleagues explicitly seek explanations for the discrepant results, but find none specifically related to the 2 studies that appear plausible. They clearly recognize the possibility that the improved EFS and OS in patients given escalated dose ara-C + G-CSF in HOVON-42 will eventually prove to be a chance observation, even though they adjusted the above-noted P values to reflect the several tests of statistical significance they performed.

Therapeutic findings aside, Pabst et al's report is an important reminder of the limitations of even very well conducted randomized trials (phase 3) such as AML-29 and HOVON-42. There are several reasons why such trials may prove misleading. Most basically, as the authors imply, the results are statistics, not facts. Assume that among 100 new treatments for AML, 90 are truly not useful while 10 are truly useful; history suggests this is not unrealistic.[5] Further assume a phase 3 trial formulated to have a 5% false positive rate (ie, $P = .05$) and a 20% false negative rate (ie, power = 80%). Eight of the 10 truly useful treatments will be called useful as will 4 of the truly not useful treatments. Hence,

33% (4/12) of the treatments called useful will be false positives. Of course, the false positive rate rises above 5% as the number of tests of statistical significance performed increases. In this connection, Tannock reported that subgroup analyses were done in 59% of 32 randomized trials published in the *New England Journal of Medicine* or the *Journal of Clinical Oncology*, with corrections for multiple testing done in only 13%.[6] Under the circumstances, it is quite plausible that more than 50% of the treatments reported as advances are not.

Various biases also need to be considered. Publication bias is well known[7] and has motivated the establishment of trial registries that include unpublished as well as published results. Although not an issue with the current article, funding source influences whether an experimental treatment will be concluded to be "treatment of choice" after a phase 3 trial. Thus, Als-Nielsen et al found that, after accounting for treatment effect, double-blinding, and other covariates, trials funded by for-profit organizations were 5.3-fold more likely to recommend the experimental treatment (95% confidence interval 2.0-14.4).[8]

A fundamental purpose of randomization is to achieve balance on unknown covariates. The latter's importance in AML is apparent given that accounting for known covariates (cytogenetics, FLT3, etc) results in an ability to predict long-term outcomes that is closer to a coin flip (area under receiver operating characteristic curve, AUC, = 0.5) than certainty (AUC = 1.0).[9] Yet randomized trials of a given therapy are conducted sequentially and thus perforce differ with respect to these unknown covariates. Hence, the results of randomized trials of the same therapy may not be mutually consistent unless the treatment effect is quite large.

A recent paper was provocatively entitled "Why most published research findings are false."[10] Even without necessarily subscribing to this view, physicians' seeming reluctance to be influenced by results of even randomized trials is understandable, even if the reasons for this reluctance are often intuitive. An adage attributed to the late legendary college basketball coach John Wooden is, "Be quick, but do not hurry." Perhaps, and depending on effect size, we should be quick to organize follow-up trials to confirm "positive" results of well-conducted trials such as that of Pabst et al, but circumspect in altering practice to reflect the results.

Conflict-of-interest disclosure: The author declares no competing financial interests. ■

REFERENCES

1. Pabst T, Vellenga E, van Putten W, et al. Favorable effect of priming with granulocyte colony-stimulating factor in remission induction of acute myeloid leukemia restricted to dose escalation of cytarabine. *Blood*. 2012; 119(23):5367-5373.

2. Uy G, Avigan D, Cortes J, et al. Safety and tolerability of plerixafor in combination with cytarabine and daunorubicin in patients with newly diagnosed acute myeloid leukemia-preliminary results from a phase I study [abstract]. *Blood*. 2011;118:82.

3. Estey E. Use of colony-stimulating factors in the treatment of acute myeloid leukemia. *Blood*. 1994; 83(8):2015-2019.

4. Lowenberg B, van Putten W, Theobald M, et al. Effect of priming with granulocyte colony-stimulating factor on the outcome of chemotherapy for acute myeloid leukemia. *N Engl J Med*. 2003;349(8):743-752.

5. Walter R, Appelbaum F, Tallman M, Weiss N, Larson R, Estey E. Shortcomings in the clinical evaluation of new drugs: acute myeloid leukemia as paradigm. *Blood*. 2010;116(14):2420-2428.

6. Tannock IF. False-positive results in clinical trials: multiple significance tests and the problem of unreported comparisons. *J Natl Cancer Inst*. 1996:88(3-4):206-207.

7. Simes R. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol*. 1986;4(10):1529-1541.

8. Als-Nielsen B, Chen W, Gluud C, Kjaergard L. Association of funding and conclusions in randomized drug trials. *JAMA*. 2003;290(7):921-928.

9. Walter R, Othus M, Borthakur G, et al. Quantitative effect of age in predicting empirically-defined treatment-related mortality and resistance in newly diagnosed AML: case against age alone as primary determinant of treatment assignment [abstract]. *Blood*. 2010;116:2191.

10. Ioannidis J. Why most published research findings are false. *PLoS Med*. 2005;2(8):0696-0701.

● ● ● RED CELLS & IRON

Comment on Burger et al, page 5512

Marker-of-self becomes marker-of-senescence

Nisha Sosale and Dennis E. Discher UNIVERSITY OF PENNSYLVANIA

In this issue of *Blood*, Burger et al provide compelling in vitro evidence that phagocytic uptake of aged human erythrocytes is promoted by the marker of self CD47 in complex with the serum factor thrombospondin-1.[1] CD47 is expressed on cells of all types and binds and signals to SIRPα on phagocytes.

The key in vivo evidence thus far for a marker-of-self role for CD47-SIRPα interactions comes from CD47-knockout mice: when red cells from these mice are injected into the circulation of control mice, the deficient cells are cleared within hours by macrophages in the spleen whereas normal red cells circulate for weeks.[2] Surprisingly, the knockout mice are not anemic, but additional in vitro studies with both mouse and human red cells[3] have largely confirmed that the "eat me not" signal can counter "eat me" signals on red cells—despite one report to the contrary.[4] Loss of CD47's protective role had long been hypothesized to impact clearance of senescent cells,[2] but data have been lacking. The studies of Burger et al suggest instead that aged cells add thrombospondin-1 to the CD47-SIRPα interaction (see figure), and this then prompts phagocytosis of red cells.

The biophysical nature of the binary CD47-SIRPα interaction must be suitably tuned to signal "self" through phosphorylation of SIRPα's inhibitory motif in the cytoplasmic tail. If the CD47-SIRPα interaction is too weak, then signaling will be insufficient. On the other hand, if the CD47-SIRPα interaction is too strong, then the macrophage cannot let go of the red cell, even when exposed to blood flow through the spleen that tends to disrupt erythrocyte-macrophage adhesion. In vitro studies of increasing CD47 density on microparticles have shown that signaling to SIRPα increases to saturation in correlation with CD47's real but limited ability to inhibit phagocytosis.[3] The culture studies of Burger et al do not examine SIRPα signaling upon addition of thrombospondin-1, but their assays do suggest that the ternary interaction is stronger than the binary because aged cells, but not fresh cells, adhere strongly to SIRPα-expressing cells. This strong adhesion evidently favors "eat me."

Living in an oxidative environment as we do, red cells accumulate oxidative damage that could include changes to CD47, according to