

Acting on Imperfect Evidence: How Much Regret Are We Ready to Accept?

Benjamin Djulbegovic *H. Lee Moffitt Cancer Center and Research Institute at the University of South Florida, Tampa, FL*
Andrew Frohlich and Charles L. Bennett, *Jesse Brown VA Medical Center and the VA Midwest Center for Health Services Research and Policy Studies, the Division of Hematology/Oncology of the Department of Medicine, the Institute for Health Services Research and Policy Studies, and the Robert H. Lurie Comprehensive Cancer Center of Northwestern University, Chicago, IL*

In this issue of the *Journal of Clinical Oncology*, Papaldo et al¹ report the findings of a large nonrandomized trial evaluating the role of prophylactic recombinant granulocyte colony-stimulating factor (filgastrim) for the prevention of febrile neutropenia in early breast cancer patients receiving relatively high-dose epirubicin plus cyclophosphamide. Previous randomized clinical trials and meta-analyses have identified the ability of filgastrim support in decreasing the duration of neutropenia and reducing the incidence of febrile neutropenia when cancer patients received intensive and/or high-intensity chemotherapeutic regimens.²⁻⁴ Recommendations from the American Society of Clinical Oncology (ASCO) support the prophylactic use of filgastrim in conjunction with chemotherapeutic regimens that are associated with 40% or higher rates of febrile neutropenia, and those from the National Comprehensive Cancer Center Network support a 20% cut point.^{5,6} The drug is one of the most commonly prescribed supportive-care agents today and has important benefits on duration of neutropenia and hospitalization for febrile neutropenia treatment.

This observational study finds that primary prophylaxis with filgastrim can decrease the frequency of severe neutropenia, febrile neutropenia, delayed chemotherapy cycles, and chemotherapy dose reductions in breast cancer patients receiving relatively high-dose chemotherapy. Of particular interest, the study found that filgastrim administration at a dose of 300 $\mu\text{g}/\text{d}$ on days 8 and 12 was associated with outcomes similar to those of more intensive filgastrim schedules. What does this study add to our current knowledge about filgastrim use, and should it lead to downstream changes in clinical research and clinical practice?

The study directly addresses the central issue of filgastrim scheduling. Whereas the initial placebo-controlled random-

ized trials evaluated daily dosing of filgastrim until absolute neutrophil counts of 10,000 cells/ mm^3 were reached following administration of intensive-dose chemotherapy regimens to persons with small-cell lung cancer, clinicians and the ASCO guidelines advocate shorter schedules of filgastrim.^{2,3} This study raises the important question as to whether even lower-dose and less frequent filgastrim schedules are likely to be useful. The study evaluated 506 women with early breast cancer who participated during 1991 to 1994 in a factorial 2 \times 2–design study of epirubicin 120 mg/m^2 and cyclophosphamide 600 mg/m^2 intravenously on day 1 every 21 days for four cycles. The experimental agents included filgastrim and lonidamine, an agent that is capable of reversing resistance to anthracyclines in vitro.⁷ Five consecutive cohorts of 100 patients each received filgastrim schedules of 480 $\mu\text{g}/\text{d}$ on days 8 to 14; 480 $\mu\text{g}/\text{d}$ on days 8, 10, 12, and 14; 300 $\mu\text{g}/\text{d}$ on days 8, 10, 12, and 14; 300 $\mu\text{g}/\text{d}$ on days 8 and 12; or no filgastrim. The very short schedules were designed to administer one dose of filgastrim before the expected neutrophil nadir time, and a second dose during the neutrophil nadir time. Given the potential clinical and economic implications for this relatively costly drug, this study provides important hypothesis-generating data suggesting that shorter courses of filgastrim may be sufficient for prophylaxis of neutropenia in this setting.

However, this is a large, nonexperimental, observational study. Therefore, one can never be sure if the observed outcomes were because of the effect of granulocyte colony-stimulating factor or because of selection bias, which can rarely be confidently ruled out in nonrandomized trials. Thus, the most important role of this study is to inform the design of subsequent large phase III clinical trials of low- versus standard-intensity filgastrim support. If the goal is to provide reliable evidence, there is no statistically

and clinically acceptable shortcut for bypassing randomization. Sequential cohorts provide insights into the potential of alternative prophylactic filgrastim schedules, though no definitive conclusions about the likely efficacy and toxicity of the various schedules can be made with this design.

Nevertheless, the study may have implications for clinical practice. It is still fair to ask the following questions: Is it acceptable to change practice based on nonrandomized clinical trial findings? What should we do when studies suggest that possibly better treatments can be recommended but definitive evidence from randomized clinical trials has not been reported?

We propose considering the following principles when evaluating these questions:

First, one should consider the overall purpose of (clinical) research. The purpose of clinical research is to address uncertainties about the effects of competing treatments: if there were no uncertainties, there would be no need to conduct research.^{8,9} Scientifically, nothing new would be learned, and ethically we would be exposing patients to risks associated with potentially inferior treatment. Uncertainty comes in various forms, ranging from simply not knowing, to a high degree of uncertainty. The choice of scientific methodology and particular research design should reflect the level of uncertainty.¹⁰ Therefore, the choice of randomized versus nonrandomized design should be viewed from the perspective of tailoring a research design to varying levels of uncertainty. For example, few would argue that the bleeding patient should be randomly

Table 1. Proposed Evidentiary Standards for Clinical Recommendations As a Function of Treatment Goals and Acceptable Regret

Goals of Treatment	Uncertainty About Benefits/Harms of Alternative Treatment Options	Benefit-Harm Ratio	Regret Associated With Wrong Recommendations	Acceptable Evidentiary Standards	Examples
Prevention of the disease in asymptomatic (healthy) individuals	High	There are important trade-offs between the benefits and harms	High	Highest standard of experimental evidence (eg, well-designed and conducted large-scale RCTs/systematic reviews based on large randomized evidence)	Tamoxifen in prevention of breast cancer
Cure	Low	The vast majority of practitioners believe that the intervention does more good than harm	Low	May accept lower level of evidence (eg, small RCTs, observational studies, phase II trials with large effect size, etc)	Antibiotics in <i>Helicobacter pylori</i> -positive MALT gastric lymphoma; surgery of isolated liver metastasis of colorectal cancer
Prolongation of survival in years	Low	The vast majority of practitioners believe that the intervention does more good than harm	Low	May accept lower level of evidence (eg, small RCTs, observational studies, phase II trials with large effect size, etc)	Imatinib in chronic myeloid leukemia; cladribine in hairy-cell leukemia
Prolongation of survival by days to months	High	There are important trade-offs between the benefits and harms, or it is not clear whether the intervention does more good than harm	High	Highest standard of experimental evidence (eg, well-designed and conducted large-scale RCTs/systematic reviews based on large randomized evidence)	Chemotherapy in metastatic lung cancer
Palliation (improvement of quality of life)	Low/moderate	The vast majority of practitioners believe that the intervention does more good than harm	Low	May accept lower quality of evidence (eg, observational studies, case series, inferential judgment) if costs are low	Morphine for pain control
Palliation (improvement of quality of life)	High	It is not clear whether the intervention does more good than harm	Moderate	High-quality evidence (eg, meta-analysis, RCT etc) particularly if costs are high	Bisphosphonates in prevention of skeletal-related morbidities; erythropoietin in treatment of chemotherapy-related anemia; granulocyte colony-stimulating factor in treatment/prevention of febrile neutropenia

Abbreviation: RCT, randomized controlled trial.

assigned to blood transfusion versus no blood transfusion to prove that transfusion is superior, though clinical trials have evaluated low versus high hemoglobin thresholds for transfusion in initiations. When the effects of treatments are dramatic and readily recognizable, we do not need randomized trials to demonstrate that a particular intervention is effective.¹¹ However, when estimated uncertainties about benefits and harms of competing treatment alternatives are high, we can never be sure if the observed findings were due to treatment(s) or because of other factors related to patient selection or biases associated with the chosen methods of study. Under these circumstances, there is simply no better method than a randomized controlled trial to resolve these uncertainties.

Second, we should be cognizant of the fact that there are always possibilities that ineffective treatments are recommended or even that effective treatments are not recommended. In other words, we may ultimately regret our decisions and recommendations. Formally, regret for not recommending effective treatments becomes acceptable when chances of successes and expected treatment benefits are low.^{12,13} Committing to treatment will make regret tolerable only if the chances of treatment success are higher than the harms associated with these therapies (which may include costs).^{12,13}

Third, our “tolerance toward error”^{8,12,13} depends in large part on the goals of specific treatments.^{12,13} When the goal is prevention of disease, such as with tamoxifen as prophylaxis against breast cancer or small improvements in expected survival durations, we should insist on the highest level of evidence. This would typically include systematic reviews based on large-scale randomized evidence. However, if the goals of treatment are cure or improvements in years of expected survival, with high perceived benefits and low or unknown extent of perceived risks, a lower level of evidence may be acceptable. If the goal of the treatment is less meaningful prolongation of survival (eg, days to months as is often seen in lung cancer chemotherapy trials), with uncertain benefits and harms, then our practice should be based on the highest evidentiary standards possible. In evaluating symptomatic treatments associated with high costs and/or quality-of-life benefits,

a high quality of evidence is needed to facilitate societal decision making.

Table 1 presents the principles outlined above and examples of how and where clinicians can consider goals of treatment, uncertainties about benefits and harms of alternative treatment, and regret associated with wrong recommendations. These factors should assist with the choice of acceptable evidentiary standards for given clinical scenarios.^{12,13}

In the context of these questions, deciding to accept non-randomized evidence for filgastrim scheduling depends on our perceived goals of treatment, expected costs and quality-of-life implications of therapy, and our tolerance for regret.^{8,12,13} These concerns are relevant in light of recent clinical trials, particularly the study that found that women with node-positive early breast cancer had significantly improved clinical outcomes (survival) with dose-dense combination chemotherapy with filgastrim support.¹⁴ Papaldo et al¹ hypothesize that four every-other-day filgastrim injections might be sufficient to support the dose-dense combination chemotherapy regimen, citing the favorable findings of this filgastrim regimen in conjunction with a biweekly chemotherapy regimen that was administered to persons with advanced gastric cancer. However, if we were to adopt the recommendations of Papaldo et al,¹ the potential exists for high regret in this setting. Therefore, a randomized clinical trial should be initiated; one that evaluates shorter versus longer schedules of filgastrim prophylaxis among breast cancer patients who receive dose-dense combination chemotherapy regimens. The study design should include formal assessments of clinical outcomes, quality of life, and direct and indirect medical costs.

In conclusion, the findings from the study by Papaldo et al¹ are unlikely to be sufficient to inform changes in clinical practice, because of concerns related to both costs and quality of life. While we agree with the recently reported warning of Browman¹⁵ that standards of proof need to be high when advocating for changes in standards of practice, there will be situations when acting on findings from non-randomized clinical trials will not cause us “chagrin.”¹⁶ However, filgastrim scheduling is unlikely to be the proper setting to challenge our tolerance for regret.

© 2005 by American Society of Clinical Oncology

Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following authors or their immediate family members indicated a financial interest. No conflict exists for drugs or devices used in a study if they are not being evaluated as part of the investigation. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Authors	Employment	Leadership	Consultant	Stock	Honoraria	Research Funds	Testimony	Other
Benjamin Djulgovic					Amgen (A)			
Charles L. Bennett		Amgen (A)				Amgen (B)		
Dollar Amount Codes (A) < \$10,000 (B) \$10,000-99,999 (C) ≥ \$100,000 (N/R) Not Required								

REFERENCES

1. Papaldo P, Lopez M, Marolla P, et al: The impact of five prophylactic filgrastim schedules on hematologic toxicity in early breast cancer patients treated with epirubicin and cyclophosphamide. *J Clin Oncol* 23:6908-6918, 2005
2. Trillet-Lenoir V, Green J, Manegold C, et al: Recombinant granulocyte colony stimulating factor reduces the infectious complications of cytotoxic chemotherapy. *Eur J Cancer* 29A:319-324, 1993
3. Crawford J, Ozer H, Stoller R, et al: Reduction by granulocyte colony-stimulating factor of fever and neutropenia induced by chemotherapy in patients with small-cell lung cancer. *N Engl J Med* 325:164-170, 1991
4. Clark OAC, Lyman GH, Castro AA, et al: Colony stimulating factors for chemotherapy-induced neutropenia: A meta-analysis. *J Clin Oncol* 23:4198-4214, 2005
5. Ozer H, Armitage JO, Bennett CL, et al: Update of recommendations for the use of hematopoietic colony stimulating factors: Evidence-based, clinical practice guidelines. *J Clin Oncol* 18:3558-3585, 2000
6. Myeloid Growth Factors in Cancer Treatment. Accessed May 31, 2005. http://www.nccn.org/professionals/physician_gls/PDF/myeloid_growth.pdf
7. Papaldo P, Lopez M, Cortesi E, et al: Addition of either lomadine or granulocyte colony stimulating factor does not improve survival in early breast cancer patients treated with high dose epirubicin and cyclophosphamide. *J Clin Oncol* 21:3462-3468, 2003
8. Djulbegovic B: Acknowledgment of uncertainty: A fundamental means to ensure scientific and ethical validity in clinical research. *Curr Oncol Rep* 3:389-395, 2001
9. Chalmers I: Well informed uncertainties about the effects of treatments. *BMJ* 328:475-476, 2004
10. Courtney H, Kirkland J, Viguierie P: Strategy under uncertainty. *Harvard Bus Rev* 1-32, 1999
11. Chalmers I: Uncertainties about the effects of treatments-how to deal with them, in Chalmers I, Thornton H, Evans I (eds): *Evaluating the Effects of Medical Treatments*. London, UK, BMJ Books
12. Djulbegovic B, Hendler F, Pavletic S: Reasoning by identifying goals of treatment. *Cancer Control* 6:377-384, 1999
13. Djulbegovic B, Hozo I, Schwartz A, et al: Acceptable regret in medical decision making. *Med Hypotheses* 53:253-259, 1999
14. Citron ML, Berry DA, Cirrincione C, et al: Randomized trial of dose-dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: First report of Intergroup Trial C9741/Cancer and Leukemia Group B Trial 9741. *J Clin Oncol* 21:1431-1439, 2003
15. Browman GP: Standards of proof, standards of practice, and proof of standards: A tale of two trials. *J Clin Oncol* 23:2583-2585, 2005
16. Feinstein AR: The 'chagrin' factor and qualitative decision analysis. *Arch Intern Med* 145:1257-1259, 1985