The Hodi Paper: A Researcher's View

CHRISTOPHER R. FRIESE, RN, PhD, AOCN®

From University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Correspondence to: Christopher R. Friese, RN, PhD, AOCN®, Assistant Professor, University of Michigan, 400 North Ingalls #4162, Ann Arbor, MI 48109-5482. E-mail: cfriese@umich.edu

J Adv Pract Oncol 2010;1:287–289

Review of: "Improved survival with ipilimumab in patients with metastatic melanoma," by Hodi et al. (2010). The New England Journal of Medicine, 363(8), 711–723. doi:10.1056/NEJMoa1003466.

The data reported by Hodi et al. (2010) present a bright spot for patients with unresectable stage III and IV melanoma. This review will address some of the methodological considerations of the study, as Ms. Esper has highlighted the clinical implications. A deeper understanding of the methods used in this clinical trial will help us better interpret the findings. I shall focus my remarks on the randomization scheme, the shift in primary endpoint, and the challenges faced when no "gold standard" therapy is available.

Is the Randomization Truly Random?

A careful read of the manuscript reveals that patients were assigned randomly in the following ratio: 3 patients in the ipilimumab-plus–glycoprotein 100 (gp100) arm, to 1 patient in the ipilimumab-alone arm, to 1 patient in the gp100 alone arm. The study also stratified the randomization scheme by two variables: baseline metastatic stage and prior receipt of interleukin-2 (IL-2, Proleukin) therapy. Because of this unequal allocation scheme and the set of criteria in which randomization occurred, a reader might question the "randomness" of the patient assignment. Let us review the approach and the possible rationale.

UNEQUAL ALLOCATION

First, let us discuss the 3:1:1 treatment allocation. In a classic experiment, subjects are randomly allocated in equal groups. This is done to maximize the statistical power to compare differing treatment groups. An important assumption in the classic approach is that the variability in outcomes is the same for each treatment. As already noted, investigators are still

searching for a gold standard induction therapy for unresectable malignant melanoma. Thus, experimental therapeutics using novel agents is an important aspect of melanoma research.

One reason to allocate unevenly is to accrue the required number of patients more quickly and gain experience. Because ipilimumab and gp100 are relatively new agents, investigators have less experience with them, and thus are less familiar with their potential side effects. Unequal allocation to the ipilimumab-plus-gp100 arm—in this case the arm of the study with the highest risk for side effects—would speed up recruitment efforts and give investigators more experience with these agents.

A second reason to allocate unevenly is to enable subgroup analysis. A higher proportion of patients in this arm also increases the ability to detect differences in patient subgroups that are statistically and clinically meaningful. This is precisely what was done in the Hodi study, where overall survival across treatments was assessed by subgroups, such as gender, age, and baseline lactate dehydrogenase. Such an analysis helps us identify subgroups that benefit from a particular regimen.

A third reason to unevenly allocate subjects is when the outcome response in one arm varies more than the other treatment arms. In this case, the combination of two novel agents is likely to have more variation in outcome than either agent by itself. An increase in patients reduces the "noise" detected in the statistical analysis for this group.

In this study, one can see why unequal allocation may be preferable to a traditional 1:1:1 approach. Using unequal allocation, empirical studies have demonstrated no loss in statistical power when more than two treatment groups are simultaneously studied. This schema has been used in many other important studies, such as the British Doctors' Trial, where patients were randomized 2:1 to receive aspirin or placebo (Peto et al., 1988). The original intent was to compare two doses of aspirin in the aspirin arm (i.e., a subgroup analysis). Although the dual-dose aspect of the study was dropped, the 2:1 allocation gave the investigators ample statistical power to examine reasons for premature discontinuation in the aspirin arm. This analysis led us to better understand the short- and long-term side effects of prophylactic aspirin therapy.

STRATIFIED RANDOMIZATION

But why stratify the randomization scheme by clinical factors? Shouldn't randomization alone balance the subgroups? There is always a chance that the randomization scheme will result in groups that are unbalanced by characteristics known to influence the outcome. In this case, clinical variables, such as stage and prior IL-2 therapy, are known to influence melanoma progression. Thus, it is appropriate to stratify the results by these variables. In the paper, Hodi and colleagues identified other prognostic variables, such as lactate dehydrogenase, age, and sex. Why not also stratify the random assignment by these variables? As the number of variables used to stratify increases, so too does the chance for groups to go empty. The clinical challenge is to select one or two of the most meaningful prognostic variables by which to stratify the randomization.

Changing Primary Endpoints: Moving the Goal Posts?

The authors stated that the trial's original primary endpoint was overall response, which was defined as patients who obtained a partial or complete response. However, after interim data were reviewed, the primary endpoint was changed to overall survival. A likely first question is, "Why not study survival initially?"

Metastatic melanoma is a disease that is historically resistant to treatment. Thus, in the setting of novel agents, it is premature to expect to detect a survival advantage in heavily pretreated patients. One benefit of the modern clinical trials system is the important function of data safety monitoring boards (DSMBs)—groups of informed reviewers detached from the study who examine interim results. In this case, promising data emerged regarding these two agents. This led the investigative team to amend their protocol, and with human subjects committee approval, examine both response and survival. If the agents showed noteworthy clinical activity, the U.S. Food and Drug Administration (FDA) might consider the data for approval earlier than previously anticipated. It is important to note that the study's primary endpoint moved from overall response to overall survival, and not the other way around. Had the reverse been true, our enthusiasm for these agents would be diminished.

Searching for the Gold Standard

In an accompanying editorial, Hwu (2010) questioned why dacarbazine (DTIC) was not included in the trial design. This is an appropriate question, as both IL-2 and dacarbazine are FDA-approved for the treatment of metastatic melanoma, and both have shown clinical activity in the disease. Although IL-2 was considered in the stratification, dacarbazine was not. This concern highlights the clinical challenge when no gold standard therapy is available.

In this case, we are left to examine the historical record of overall response and survival when patients are treated with IL-2 and/or dacarbazine. It is likely that this question will be asked when the FDA reviews the data, and a new study to examine ipilimumab to IL-2 or dacarbazine may be warranted, even though these therapies are admittedly suboptimal.

Conclusions

This study has important implications for patients with metastatic melanoma, and the providers who care for them. Because of a deliberate unequal allocation and stratified randomization, the authors gained more experience with these novel agents, compared clinically important subgroups, and minimized the chance of variable results without a loss in statistical efficiency. The change in primary endpoints through the mechanism of the DSMB was a sound strategy and lends credibility to the study's results. Until further studies comparing these active agents are conducted, we as clinicians are left with the difficult question of what to do with the patients we see in clinic tomorrow.

DISCLOSURES

The author has no potential conflicts of interest to disclose.

REFERENCES

Hodi, F. S., O'Day, S. J., McDermott, D. F., Weber, R. W., Sosman, J. A., Haanen, J. B., . . . Urba, W. J. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *The New England Journal of Medicine, 363*(8), 711–723. doi:10.1056/NEJMoa1003466

Hwu, P. (2010). Treating cancer by targeting the immune system. *New England Journal of Medicine, 363(8),* 779–781.

Peto, R., Gray, R., Collins, R., Wheatley, K., Hennekens, C., Jamrozik, K., . . . Norton, S. (1988). Randomised trial of prophylactic daily aspirin in British male doctors. *British Medical Journal (Clinical Research Ed.), 296*(6618), 313–316. doi: 10.1136/bmj.296.6618.313