Letter to the Editor

Re: Impact of the First Generation of Children’s Oncology Group Clinical Trials on Clinical Practice for Wilms Tumor

I read with interest the article titled, “Impact of the First Generation of Children’s Oncology Group Clinical Trials on Clinical Practice for Wilms Tumor” by Dome et al., which provides an excellent summary of some of the results from the recent clinical trials for children with favorable or anaplastic histology Wilms tumor conducted by the Children’s Oncology Group (COG). However, the article discusses a treatment approach that they identify as “a good standard of care,” a claim they suggest is justified by the results of the studies they have summarized. The results of 2 of their published studies require careful review and I assert do not establish a good standard of care.

Dome et al. suggest that regimen M is a good standard of care for patients with favorable-histology Wilms tumor (FHWT) and pulmonary metastases limited to the lungs that fail to resolve completely after an initial 5-week period of chemotherapy with the combination of vincristine/dactinomycin/doxorubicin, based on a comparison of the event-free and overall survival of patients with FHWT and metastases limited to the lungs treated on AREN0533 versus the treatment outcomes for a control group of patients treated on National Wilms Tumor Study (NWTS)-5. There are, however, details regarding the control group that render the comparison invalid.

Patients were eligible for AREN0533 if they had pulmonary metastases identified on a chest CT. The comparison group was defined by the presence of pulmonary metastases on a chest radiograph or chest CT and treatment with whole-lung radiation therapy (WLRT). Thus, patients with pulmonary metastases who were not treated with WLRT on NWTS-5 were considered by the COG investigators to not have stage IV disease and were not included in the comparison group. Only 42% of patients enrolled on NWTS-5 who had pulmonary nodules were treated with WLRT. Thus, the NWTS-5 comparison group of patients used by the COG investigators was clinically different from the patients enrolled on AREN0533.

The COG investigators did not state how many potentially eligible patients enrolled on AREN03B2 during the period when AREN0533 was open for accrual were actually enrolled on AREN0533. In another study conducted by the group that examined augmentation of treatment for patients with FHWT and combined loss of heterozygosity (LOH) of 1p and 16q, only 65.3% (32/49) of eligible patients with stage I or II FHWT and 62.2% (51/82) of eligible patients with stage III or IV FHWT were enrolled. No comparison of the differences between the eligible and enrolled populations was published. It is not clear from the published data what percentage of potentially eligible patients was enrolled on AREN0533.

The COG investigators reported that the comparison group from NWTS-5 included 268 patients with stage IV FHWT and pulmonary metastases only. NWTS-5 enrolled only 242 patients with stage IV FHWT in the “studied” category (biology and clinical data). If 14.9% (the percentage with extrapulmonary metastases enrolled on AREN0533) had extrapulmonary metastases, there would have been only 204 available for comparison, not 268. If all but those NWTS-5 patients who were included in the “registered only” category are included in the comparison group, the total with stage IV FHWT would be 310 (S.M. Peterson, 2021, unpublished data). Excluding those with extrapulmonary metastases produces a total of 264. The reason for the discrepancy is not clear, but a previous publication from the COG Renal Tumor Committee required correction because the initial published version included patients with both favorable and other histologies when only favorable histology patients were to be included in the comparison group. The revised comparisons to NWTS-5 were no longer statistically significant.

Evaluation of the effectiveness of regimen M for patients whose pulmonary nodules do not respond completely to 6 weeks of 3-drug chemotherapy is limited by the removal from the analysis of 2 patients who developed progressive disease prior to completion of 4 cycles of regimen M chemotherapy and WLRT. The COG investigators did not indicate that they had removed patients who developed progressive disease prior to 3 months after diagnosis from the NWTS-5 comparison group.

The COG investigators state that, “Augmentation of therapy for patients with FHWT and combined LOH 1p/16q reduces the risk for relapse.” However, as noted earlier, not all eligible patients were enrolled in the study. Because the COG investigators defined stage IV based on the presence of pulmonary nodules only, whereas they defined stage IV for the NWTS-5 patients as the presence of pulmonary nodules and treatment with WLRT, the stage distribution of those enrolled on AREN0532 and AREN0533 was very different from that of the comparison sample from NWTS-5. AREN0522 appeared to have substantially more patients with stage IV disease (24.1% vs 5.9%), whereas NWTS-5 appeared to have more patients with stage I (17.6% vs 9.6%) and stage II disease (38.8% vs 28.9%). Augmented therapy did not produce a statistically significant improvement in overall survival compared with NWTS-5, and did expose all patients with stage I and II disease to doxorubicin, rather than only the 31.2% who would have received it as relapse treatment, and exposed all patients with stage III and IV disease to etoposide and cyclophosphamide, which would have been included in a relapse treatment regimen for only the 39.7% who experienced relapse. In the absence of a comparison between all of those enrolled in AREN03B2 who had FHWT and combined LOH at 1p and 16q versus those actually enrolled on AREN0532 and AREN0533 with the same characteristics, and modification of the definition of stage IV for the NWTS-5 patients used for comparisons with AREN0533 patients with stage IV FHWT, suggestions that augmentation of therapy benefits patients with FHWT and combined LOH at 1p and 16q are not justified.

The phrase “good standard of care” does not imply that the recommendation is based on high-quality data, such as from a randomized clinical trial or a confirmatory single-arm trial. Given the significant shortcomings of the studies reported by the COG investigators, I believe the approaches outlined in the 2 published manuscripts should not be interpreted as a standard of care, but as results of 2 studies with
significant methodological flaws that require replication in appropriately designed, randomized studies.

Daniel M. Green, MD
Departments of Oncology, and Epidemiology and Cancer Control,
St. Jude Children’s Research Hospital,
Memphis, Tennessee
Email: daniel.green@stjude.org

References


Authors’ Reply

Reply to the Letter to the Editor by Daniel M. Green

We are pleased to provide our perspectives on the points raised by Dr. Green regarding our article, “Impact of the First Generation of Children’s Oncology Group Clinical Trials on Clinical Practice for Wilms Tumor.” Dr. Green has previously raised similar points in letters to the editor, to which we have responded.

The first point regards our conclusion that regimen M represents a good standard of care for patients with favorable-histology Wilms tumor (FWHT) and isolated lung metastases that do not respond completely to 6 weeks of chemotherapy. The AREN0533 study was designed to improve event-free survival (EFS) from 75% to 85% for patients with incomplete lung nodule response by adding cyclophosphamide/etoposide to the standard vincristine/dactinomycin/doxorubicin combination (regimen M). These EFS targets were based on outcomes according to attainment of complete response at day 70 for patients with isolated lung metastases treated on the fifth National Wilms Tumor Study (NWTS-5). Under the null hypothesis EFS of 75%, the expected event rate was 25% (30 of 120 subjects), but the observed event rate was only 12.2% (16 of 131 subjects; P = .001), leading to the firm conclusion that regimen M is beneficial for this patient group.

In a separate analysis to which Dr. Green referred, we compared the outcomes of patients with FWHT and isolated pulmonary metastases treated on AREN0533 to those treated on NWTS-5. This analysis was undertaken in response to an external review suggesting that we evaluate the totality of changes enacted for patients with isolated lung nodules on AREN0533, which included therapy reduction (omission of lung radiation therapy [RT]) for those with complete lung nodule response and therapy augmentation (regimen M) for those with incomplete lung nodule response. On NWTS-5, patients with lung nodules visible on CT scan and not chest radiograph were treated according to physician choice such that some were treated as stage IV and others were treated according to the abdominal stage. For the AREN0533 comparison, we deliberately included only patients on NWTS-5 treated as stage IV with doxorubicin and lung RT. The rationale for this was that on NWTS-4 and NWTS-5, 5-year EFS for patients with CT-only nodules who received vincristine/dactinomycin/doxorubicin was 80% compared with 56% for patients who received only vincristine/dactinomycin with or without lung RT (P = .004). Likewise, 5-year EFS was 82% in patients who received lung RT and 72% in patients who did not (P = .13). Had we included all NWTS-5 patients with lung nodules regardless of treatment, the superior EFS observed with AREN0533 would likely have been magnified.

Further regarding the comparison between AREN0533 and NWTS-5, Dr. Green points out that the number of NWTS-5 patients included did not match the number he expected based on previous published analyses. The comparison involved a new analytic dataset from NWTS-5 that met the criteria of stage IV FWHT, isolated pulmonary metastasis, and treatment as stage IV disease. We confirm that “studied” and “registered” patients were included. We do not have access to the original NWTS-5 analytic dataset to account for the 4-patient difference that Dr. Green notes, but believe that this small number is unlikely to change our conclusions.

Dr. Green raises concern that a subset of patients who enrolled on the AREN03B2 Renal Tumor Biology and Classification study did not enroll on the therapeutic studies. In contrast to NWTS-5, which was a combined therapeutic and biology study, the COG studies separated AREN03B2 from the therapeutic studies. The lack of enrollment on therapeutic studies occurred for a variety of reasons, including (1) the institution did not have the therapeutic study open, (2) the institution was unable to meet the desired timing for enrollment on the therapeutic study, or (3) the patient or their clinician declined participation in the therapeutic study. Although we do not have data on the reasons for non-enrollment, we do not believe that patients who enrolled on treatment studies had clinically or biologically different disease from those who enrolled solely on the biology study. In support of this premise, recent analyses conducted in preparation for the upcoming COG AREN2231 study indicated that patients enrolled on AREN03B2 had similar outcomes to those enrolled on the contemporaneous therapeutic studies. This was true within individual disease stages and within subgroups that received the same treatment. Based on these analyses, the statistical design for the upcoming AREN2231 study will use pooled data from AREN03B2 and the therapeutic studies to generate the historical control.