The taxanes are among the most active classes of cytotoxic agents for the treatment of breast cancer and other solid tumors. They have a molecular target (the microtubule), a mechanism of action (enhanced microtubule stability), and reasonably consistent efficacy in metastatic breast cancer (1–3). After extensive testing in the metastatic setting, we routinely use three drugs, numerous combinations, and several standard dose and schedule alternatives.

Paclitaxel administered intravenously at a dose of 175 mg/m² for 3 hours every third week emerged, first, from phase 3 trials in the metastatic setting and was tested as an add-on to standard four-cycle doxorubicin and cyclophosphamide (AC) in the Cancer and Leukemia Group B (CALGB) 9344 and National Surgical Adjuvant Breast and Bowel Project (NSABP) B-28 trials with positive results (4,5). These trials were criticized for failure to use a “best” anthracycline regimen and uneven treatment duration across the arms (ie, four vs eight treatment cycles) (6). Subsequent trials that controlled the number of treatment cycles demonstrated that docetaxel given in place of 5-fluorouracil as part of a concurrent combination with AC or in place of several cycles of fluorouracil, epirubicin, and cyclophosphamide (FEC) was superior; since then, a number of trials have been reported with mostly positive results regardless of anthracycline comparator, taxane choice, dose, or schedule (7,8). A recent meta-analysis of the published studies reported relatively constant risk reductions for both recurrence and death across the two tested taxanes (9).

Despite these seemingly consistent data, controversy continues. With the goal of maximizing benefit, we ask is there a best taxane and/or a best schedule? And with a goal of limiting toxicity, we ask...
how can we limit treatment (dose size and cycle number) and/or who can we not treat?

A best taxane might be defined through consideration of activity, toxicity, and expense, but these many variables complicate interpretation of the available head-to-head trials of the three agents in the metastatic setting, which have not established a consistent and overall drug-specific advantage that is independent of dose and schedule (10).

The optimal dose and schedule is probably agent specific. Weekly paclitaxel appears to be consistently more active and differently toxic (11–13). For docetaxel, weekly administration does not offer the same advantage, and it is worth noting that some patients and clinicians would probably prefer less frequent therapy, all else being equal (14). For albumin-bound paclitaxel, phase 3 comparative data are not yet available on this issue (15,16).

Patient selection for adjuvant chemotherapy is a key goal of current research programs. In the metastatic setting, there are (or perhaps must be) some tumors that are simply not sensitive to chemotherapy or not sensitive to taxane therapy or that possess differential sensitivity to these agents. Notably, there are no established predictive factors used in the metastatic setting to select for or against the routine use of taxanes and no consistent preclinical data in this regard (17–19). Further, neither hormone receptor status nor HER2 expression has consistently predicted the benefits of taxanes in the metastatic setting. The adjuvant setting adds a layer of complexity because it includes some or many postoperative patients who are not destined to experience recurrence before death from other causes. Judging the benefits of specific drugs against specific tumors (ie, assessing a predictive factor) is impossible if there is no tumor to assess. In such a situation, so-called predictive factors incorporate both drug–tumor interactions and tumor–host interactions to be useful.

The burning question then is whether hormone receptor status and HER2 expression can serve as predictive factors for the use of taxanes. An unplanned retrospective analysis of the CALGB 9344 trial that was based on estrogen receptor (ER) status indicated that the benefit of adding paclitaxel was substantially greater in patients with ER-negative tumors than in patients with ER-positive tumors (4). A benefit for patients with ER-positive tumors was not, however, excluded. A subsequent analysis (20) of several CALGB trials demonstrated a consistent effect: so-called better chemotherapy was most clearly superior in patients with ER-negative tumors and the impact among those with ER-positive breast cancer was about half as large. From a practical point of view, there were potential limits on the clinical application of this result because there was no untreated control group (meaning the series does not address the “any vs none” question) and ER testing was not centrally performed or reviewed.

Recent developments challenge this CALGB hypothesis. In the subset of patients with HER2-positive breast cancer, there was some evidence that the use of targeted agents (eg, trastuzumab) levels this playing field so that, in terms of efficacy, specific chemotherapy agent choice is less critical (11,21–24). However, a recent reanalysis of data from the CALGB 9344 trial (25) demonstrated that HER2 status predicted the benefit of adding paclitaxel regardless of ER status. This latter point bears emphasis because it plainly highlights the fact that ER status in isolation cannot serve as a decision point for the use of taxanes (or any specific chemotherapy agent or regimen). In addition, other groups of investigators have not consistently identified ER status as a predictor of taxane benefits in general or with regard to schedule (9,26).

The study by Martín et al. (27) in this issue of the Journal is an important addition to the ongoing discussion of the role of the taxanes in the adjuvant setting. This trial of the Grupo Español para la Investigación del Cáncer de Mama (Spanish Group for the Investigation of Breast Cancer), GEICAM 9906, compares a better anthracycline-containing regimen (six cycles of FEC using epirubicin at a dose of 90 mg/m²) against a sequence of the same version of FEC for four cycles followed by eight doses of weekly paclitaxel. This trial directly addresses the concerns of earlier critics that an inferior and shorter anthracycline regimen loaded the dice in favor of the taxane arms. Moreover, the paclitaxel dose and schedule do appear to be the best ones as well, as demonstrated by recently reported randomized studies in both the metastatic and adjuvant settings (11,12,26). Martín et al. (27) report an overall benefit that was very consistent with the results of earlier trials. However, when they used centralized testing on a large subset (74%) of the patients, they failed to confirm the results of the earlier ER- and HER2-driven retrospective subset analyses. Simply stated, Martín et al. (27) show a benefit for 8 weeks of paclitaxel, which replaces two cycles of FEC, regardless of ER or HER2 status. This result is generally consistent with those of the Breast Cancer International Research Group 001 (7), Programmes d'Actions Concertées Sein 01 (8), and Eastern Cooperative Oncology Group E-1199 (26) trials.

For fans of biologically based subset analyses who believe that they can provide information for clinical practice that are based on exploratory studies, these data are a problem. If earlier reports led you to conclude that ER and HER2 status predict the benefits of adding taxanes, how do you explain the results of the GEICAM 9906 trial and other studies? Is weekly paclitaxel so much better than every third week treatment that it compensates for the latter’s inferior efficacy with ER-positive, HER2-negative disease? This conclusion seems unlikely given the inconsistent associations across taxanes, doses, and schedules. Was their study underpowered to exclude a difference by hormone receptor and HER2 status? This conclusion is possible, but, if we saw only their present data, we would not be tempted to pursue this question. Could laboratory testing variability explain the seemingly inconsistent results? This question is harder to address, but, notwithstanding the well-documented issues of laboratory testing variability, most laboratories get most hormone receptor and HER2 test results right. Is FEC somehow different from AC with regard to an impact on subsequent taxane benefit in ER and HER2 subsets? The recent E-1199 results suggest that it is not; however, upcoming results from the NSABP B-30 trial may allow us to address this issue (26). Finally, could hormone receptor–positive breast cancer really be a collection of more or less sensitive subtypes (ie, luminal A vs B) with less or more chemotherapy sensitivity, respectively? If individual trials varied in their randomization of these subsets, we might see the present mixture of positive and negative associations.

From a practical point of view, Martin et al. (27) underscore the clinically significant risks physicians face when using unplanned subset analyses to guide clinical practice. We need to be reminded
from time to time that hypothesis-generating subset analyses are just that—hypothesis generating—and not practice changing unless and until they are confirmed prospectively (28). In this case, the retrospectively generated hypothesis was that hormone receptor and HER2 status would predict the value of adding paclitaxel with sufficient accuracy as to allow clinicians to select patients for this chemotherapy agent. The prospective study by Martin et al. (27) join a rapidly growing list of studies that do not support that hypothesis (9, 26, 27). If you accept that the taxanes are effective, hormone receptor and HER2 status should not routinely guide your selection of patients for this therapy.

References