The Oncologist®

Commentary

Randomization and Statistical Power: Paramount in Trial Reproducibility (Even for Rare Cancers)

LAURENCE H. BAKER,[a] JOHN J. CROWLEY,[b] ROBERT G. MAKI[c]

[a]Departments of Internal Medicine and Pharmacology, University of Michigan, Ann Arbor, Michigan, USA; [b]Cancer Research and Biostatistics, Administration, Seattle, Washington, USA; [c]Departments of Medicine, Pediatrics, and Orthopaedics, Mount Sinai School of Medicine, New York, New York, USA

Disclosures: Laurence H. Baker: BioMed Valley, CytRx (C/A); **Robert G. Maki:** Eli Lilly (C/A, RF), Sanofi-Aventis (C/A). The other author indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board

Research in medicine and biology is not dictated by the laws of physics. The location of a celestial body can be determined with great accuracy with relatively simple series of equations, but multiple consistent observations are needed in medicine and biology to achieve what is considered a standard practice or a biological inference. Probability and statistics are the engines that guide these inferences. There is no escaping statistics, be it in the identification of the Higgs boson [1] or the identification of one treatment as better than another.

Examples in the recent literature underscore the loss of focus on this most basic issue in the sciences in our rush to find new treatments and to further academic careers through publications. The American Society of Clinical Oncology (ASCO) published a list of the 58 studies reported in 2011 that significantly altered the way cancer is understood or had a direct effect on patient care. Of those citations, 33 were abstracts, 29 were reported at an ASCO meeting, and 9 were announcements or press releases from the U.S. Food and Drug Administration [2]. Should an abstract or a press release serve as sufficient evidence for clinical decision-making? Arguably not.

In a fascinating study of preclinical cancer research, Begley and Ellis reported in a commentary that only 6 of 53 publications cited as landmark publications could be replicated in an independent laboratory [3]. Should a single manuscript serve as sufficient evidence to draw a scientific conclusion? Again, arguably not. In this spirit, this commentary is offered on the manuscript of Pautier et al., which compares the benefit of gemcitabine alone versus the doublet of gemcitabine and docetaxel in patients with leiomyosarcoma (LMS) [4].

The interest in gemcitabine-docetaxel in soft-tissue sarcoma stems from a phase II clinical trial led by Martee Hensley, reporting that 18 of 34 patients with metastatic leiomyosarcoma (mostly uterine) achieved a Response Evaluation Criteria In Solid Tumors partial response as best result [5]. Later, Leu et al. reported the synergy of gemcitabine and docetaxel, offering a possible biological explanation for the clinical report of Hensley of two "inactive" drugs [6–9] becoming effective when combined in a specific manner [10]. Written as a phase III study, the Sarcoma Alliance for Research through Collaboration (SARC) study 002 demonstrated that gemcitabine and docetaxel were superior in terms of progression-free overall survival compared to gemcitabine alone in patients with metastatic soft-tissue sarcoma [11].

The new publication of the TAXOGEM study from Pautier et al. arrives at different conclusions than SARC002 [4]. The main conclusions are as follows:

1. Both regimens were efficacious, and gemcitabine alone yielded comparable results with less toxicity.
2. LMS in uterus responds differently (and better) than non-uterine LMS.
3. Durable stable disease is an important endpoint for patients with LMS (progression-free survival >40%).

The use of gemcitabine-docetaxel in sarcoma was last discussed in *The Oncologist* in 2007 [12]. We have updated tables from that review that cite response and survival data of gemcitabine alone and/or in combination with docetaxel (see Table 1). In Table 2, we offer a literature review of these two thera-

Correspondence: Laurence H. Baker, D.O., Department of Internal Medicine, Division of Hematology/Oncology, University of Michigan, 24 Frank Lloyd Wright Drive, Suite A3400, P.O. Box 483, Ann Arbor, Michigan 48106, USA. Telephone: 734-998-7130; Fax: 734-998-7118; e-mail: bakerl@med.umich.edu Received May 16, 2012; accepted for publication July 27, 2012; first published online in ***The Oncologist*** *Express* on August 22, 2012. 1083-7159/2012/$20.00/0 http://dx.doi.org/10.1634/theoncologist.2012-0209

Table 1. Overall response rates and overall survival times for gemcitabine alone and/or in combination with docetaxel in metastatic/unresectable soft tissue and bone sarcomas

	Overall response rate (complete remission + partial remission)		Median overall survival (months)		
Study	**Gemcitabine**	**Gemcitabine + docetaxel**	**Gemcitabine**	**Gemcitabine + docetaxel**	**Remarks**
Hensley et al. [5]		18/34 (53%): 16/29 uterine LMS; 2/5 LMS nos		NA	
Leu et al. [10]		15/35 (43%): 12/29 (41%) without bone; 7/12 LMS nos; 3/4 angiosarcoma; 1/2 MPNST; 1/3 MFH; 3/6 bone		13	
Bay et al. [13]		21/114 (18%): 11/39 LMS nos; 5/27 uterine LMS; 5/48 other histologic subtypes nos		12.1	No clear statistical differences between LMS and other histologic subtypes; no difference between uterine STS verus others
Maki et al. [11]	4/49 (8%): 1/8 LMS nos; 0/1 uterine; 2/8 MFH/HGUPS; 1/32 other sarcoma histology	12/73 (16%): 2/9 uterine LMS; 3/19 LMS nos; 4/11 MFH, 2/8 pleomorphic liposarcoma, 1/1 RMS; 0/24 other sarcoma	11.5	17.9	Uterine LMS: G+D 2/9 PR; G arm (1 PD); Other LMS: G+D 3/19 PR (2 retroperitoneal; 1 colon); G 1/8 (abdomen)
Pautier et al. [4]	7/43 (16%): 4/21 uterine LMS; 3/22 LMS nos	6/40 (15%): 5/21 uterine LMS; 1/19 LMS nos	17.5	18	
Amodio et al. [7]	1/18 (5.5%): 1 MFH (18 STS nos)		NA		
Merimsky et al. [8]	1/18 (5.5%): 1/1 uterine LMS		NA		0/1 LMS nos; 0/2 MFH; 0/1 angiosarcoma; 0/2 liposarcoma; 0/1 ASPS; 0/10 bone
Späth-Schwalbe el al. [14]	2/18 (11%): 2/3 uterine LMS		8		0/1 LMS nos; 0/5 MFH; 0/2 liposarcoma; 0/4 MPNST; 0/1 other STS
Patel et al. [9]	7/39 (18%): 3 uterine LMS; 1 LMS nos; 1/7 MFH; 1/3 angiosarcoma; 1/9 unclassified; 0/6 miscellaneous; 0/4 liposarcoma		13.9		Note: 4/10 responders in nongastrointestinal LMS (not further defined)
Okuno et al. [15]	1/29 (4%):1/2 uterine LMS		NA		0/16 LMS nos; 0/2 MFH; 0/2 osteosarcoma; 0/1 angiosarcoma; 0/1 liposarcoma; 0/4 other STS; 0/1 sarcoma nos
Svancárová et al. [16]	1/31 (3%): 1/12 LMS nos		8.8		0/2 MPNST; 0/1 angio; 0/1 RMS; 0/15 other STS
Okuno et al. [17]	1/25 (4%): 1/1 epitheloid sarcoma		15		0/12 LMS nos; 0/4 bone; 0/8 other
Look et al. [18]	9/42 (21%): uterine LMS		NA		All 42 patients had uterine LMS
Hartmann et al. [19]	1/15 (6%): 1/15 STS nos		6		MFH 6; LMS nos 3; RMS 1; MPNST 1; hemangioendothelioma 2; sarcoma nos 2
Von Burton et al. [20]	3/46 (7%): 1/8 MFH; 1/10 LMS nos; 1/12 sarcoma nos		6		0/7 fibrosarcoma; 0/4 liposarcoma; 0/4 other STS; 0/1 bone
Wagner-Bohn et al. [21]	0/20 (0%)		NA		All pediatric: RMS (8); Ewing (4); osteosarcoma (2); neuroblastoma (3); hepatoblastoma (2); nephroblastoma (1)
Ferraresi et al. [22]	1/14 (7%): 1/1 uterine LMS		11.8		0/6 LMS nos; 0/1 MHS; 0/1 undifferentiated; 0/5 other STS
Merimsky et al. [23]	1/13 (8%): 1 LMS nos		NA		Denominators of evaluable patients unknown
Maurel et al. [24]	0/7 (0%)		NA		4 STS nos; 3 bone
Samuels et al. [25]	1/9 (11%): 1 angiosarcoma		NA		All STS; histology subtypes denominators unknown
Subtotal	41/436 (9.4%)	69/290 (24%) without Leu bone	11.6 (n = 271)	14.7 (n = 262)	
Fox et al. [26]		5/53 (9%)		NA	All bone

Abbreviations: ASPS, alveolar soft part sarcoma; G, gemcitabine; G+D, gemcitabine + docetaxel; HGUPS, high-grade undifferentiated pleomorphic sarcoma; LMS, leiomyosarcoma; MFH, malignant fibrous histiocytoma; MPNST, malignant peripheral nerve sheath tumor; NA, not available; nos, not otherwise specified; PD, progressive disease; RMS, rhabdomyosarcoma; STS, soft tissue sarcoma.

Table 2. Histologic subtypes of responders

Histologic subtype	Gemcitabine ($n = 41$)[a]	Gemcitabine + docetaxel ($n = 69$)
Leiomyosarcoma (uterine)	18/71 (25%); there are 3 additional responders but *n* is unknown	28/86 (33%)
Leiomyosarcoma (other)	6/88 (7%); there are 2 additional responders but *n* is unknown	24/94 (26%)
MFH	4/33 (12%); there is 1 additional responder but *n* is unknown; there is a study with $n = 6$, but no numerator	5/14 (36%)
Angiosarcoma	1/6 (17%); there is 1 additional responder, but *n* is unknown	3/4 (75%)
Sarcoma NOS	4/85 (5%)	5/72 (7%)
Epitheloid	1/1 (100%)	0
MPNST	0/6	1/2 (50%)
Pleomorphic liposarcoma	0/?	2/8 (25%)
Liposarcoma	0/9	
Rhabdomyosarcoma	0/9	1/1 (100%)
Subtotal	41 (100%)	69 (100%)
Bone	0/25	8/59 (14%)

[a]Some reported studies were not sufficiently detailed to permit inclusion of denominator information.
Abbreviations: MFH, malignant fibrous histiocytoma; MPNST, malignant peripheral nerve sheath tumor; NOS, not otherwise specified.

pies according to histologic type, and in the case of leiomyosarcoma, primary site of disease, examining both sarcomas of soft tissue and bone. These data suggest that the combination of gemcitabine and docetaxel has significant activity and provide context for the present study; in randomized and nonrandomized trials alike, the activity of the combination appears greater than single-agent gemcitabine.

The TAXOGEM study also used a creative design, consisting of two parallel randomized phase II trials. The text identified Simon as a source for the trial design, albeit without a specific reference. As a metric to declare a significant difference in the study, a "selection" paradigm based on response was used, not a comparative paradigm. Thus, the probabilities the study cites do not represent the power of the study, but rather the probability that the arm with the higher true response rate comes out on top at the end of the trial, a weaker comparative design. This weaker design should only be used (if ever) to choose which among competing experimental regimens deserves further testing, not to declare superiority, equivalence, or noninferiority. To be true to this spirit would mean the authors should conclude that gemcitabine-docetaxel is the winner in the uterine arm (and vice versa).

The true power calculation of this study depends on the type I error chosen, which was not included in the study design. Assuming a liberal one-sided alpha of 20%, which is favored by the National Cancer Institute, yields a power for the uterine stratum (40% vs. 50% overall response rate) of 30%, far below the 80% power used in most screening studies. For the nonuterine stratum (20% vs. 40% overall response rate), the power is 55%. In other words, the parallel phase II studies are not powered to make a determination of which treatment is better. Although randomization is important, without sufficient power, the data of this study cannot be used to judge the relative benefit of gemcitabine and docetaxel versus gemcitabine alone in metastatic leiomyosarcoma.

It is also important to note a common potential bias in studies using progression-free survival study designs. If the more toxic treatment leads to slower and later assessment of disease, the more toxic therapy will appear to be better just on the basis of repeating evaluations (scans) less frequently than in the standard arm. In this study, images were repeated at 6-week intervals in one group and 8-week intervals in another, which clouds interpretation of the data as well.

We would also like to correct what appear to be misconceptions of the statistical design of SARC002. The discussion of the TAXOGEM study expresses a concern for the Bayesian study design of SARC002. We concur that the frequentist foundation of statistics is far more common and even conventional as described in the TAXOGEM study. The authors indicated that the SARC study was imbalanced for LMS histology. However, this is the precise intention of an outcome-adaptive Bayesian randomized clinical trial design [27]. SARC chose a Bayesian-based "play the winner" strategy for several reasons:

1. It was hoped that such a design would reduce the number of subjects treated with the inferior regimen [28, 29].
2. The above-mentioned phase II study was successful in reducing the number of subjects needed in each stratum.

3. We wanted to explore new ways to make improvement in a rare set of diseases occur more quickly.

We applaud the efforts of the French Sarcoma Group to address the question of therapy for rare diagnoses such as metastatic leiomyosarcoma. As noted in the first part of this commentary, reproducibility is the only way we can make firm treatment recommendations. However, the data that contribute to those recommendations must be of sufficient quality, even in rare diseases, to use as guides to treatment. The lack of statistical power is as important as a concern as randomization itself in drawing conclusions about the quality of data we examine.

REFERENCES

1. Brumfiel G. Physicists declare victory in Higgs hunt. Available at http://www.nature.com/news/physicists-declare-victory-in-higgs-hunt-1.10940. Accessed August 3, 2012.

2. Vogelzang NJ, Benowitz SI, Adams S et al. Clinical cancer advances 2011: Annual report on progress against cancer from the American Society of Clinical Oncology. J Clin Oncol 2012;30:88–109.

3. Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature 2012; 483:531–533.

4. Pautier P, Floquet A, Penel N et al. Randomized multicenter and stratified phase II study of gemcitabine alone versus gemcitabine and docetaxel in patients with metastatic or relapsed leiomyosarcomas : A Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC) French sarcoma group study (TAXOGEM study). ***The Oncologist*** 2012;17:1213–1220.

5. Hensley ML, Maki R, Venkatraman E et al. Gemcitabine and docetaxel in patients with unresectable leiomyosarcoma: Results of a phase II trial. J Clin Oncol 2002;20:2824–2831.

6. Verweij J, Lee SM, Ruka W et al. Randomized phase II study of docetaxel versus doxorubicin in first- and second-line chemotherapy for locally advanced or metastatic soft tissue sarcomas in adults: A study of the European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group. J Clin Oncol 2000;18:2081–2086.

7. Amodio A, Carpano S, Manfredi C et al. Gemcitabine in advanced stage soft tissue sarcoma: A phase II study. Clin Ter 1999;150:17–20.

8. Merimsky O, Meller I, Flusser G et al. Gemcitabine in soft tissue or bone sarcoma resistant to standard chemotherapy: A phase II study. Cancer Chemother Pharmacol 2000;45:177–181.

9. Patel SR, Gandhi V, Jenkins J et al. Phase II clinical investigation of gemcitabine in advanced soft tissue sarcomas and window evaluation of dose rate on gemcitabine triphosphate accumulation. J Clin Oncol 2001;19: 3483–3489.

10. Leu KM, Ostruszka LJ, Shewach D et al. Laboratory and clinical evidence of synergistic cytotoxicity of sequential treatment with gemcitabine followed by docetaxel in the treatment of sarcoma. J Clin Oncol 2004; 22:1706–1712.

11. Maki RG, Wathen JK, Patel SR et al. Randomized phase II study of gemcitabine and docetaxel compared with gemcitabine alone in patients with metastatic soft tissue sarcomas: Results of sarcoma alliance for research through collaboration study 002. J Clin Oncol 2007;25: 2755–2763.

12. Maki RG. Gemcitabine and docetaxel in metastatic sarcoma: Past, present, and future. ***The Oncologist*** 2007; 12:999–1006.

13. Bay JO, Ray-Coquard I, Fayette J et al. Docetaxel and gemcitabine combination in 133 advanced soft-tissue sarcomas: A retrospective analysis. Int J Cancer 2006;119:706–711.

14. Späth-Schwalbe E, Genvresse I, Koschuth A et al. Phase II trial of gemcitabine in patients with pretreated advanced soft tissue sarcomas. Anticancer Drugs 2000; 11:325–329.

15. Okuno S, Edmonson J, Mahoney M et al. Phase II trial of gemcitabine in advanced sarcomas. Cancer 2002; 94:3225–3229.

16. Svancárová L, Blay JY, Judson IR et al. Gemcitabine in advanced adult soft-tissue sarcomas. A phase II study of the EORTC Soft Tissue and Bone Sarcoma Group. Eur J Cancer 2002;38:556–559.

17. Okuno S, Ryan LM, Edmonson JH et al. Phase II trial of gemcitabine in patients with advanced sarcomas (E1797): A trial of the Eastern Cooperative Oncology Group. Cancer 2003;97;1969–1973.

18. Look KY, Sandler A, Blessing JA et al. Phase II trial of gemcitabine as second-line chemotherapy of uterine leiomyosarcoma: A Gynecologic Oncology Group (GOG) study. Gynecol Oncol 2004;92:644–647.

19. Hartmann JT, Oechsle K, Huober J et al. An open label, non-comparative phase II study of gemcitabine as salvage treatment for patients with pre-treated adult type soft tissue sarcoma. Invest New Drugs 2006;24:249–253.

20. Von Burton G, Rankin C, Zalupski MM et al. Phase II trial of gemcitabine as first line chemotherapy in patients with metastatic or unresectable soft tissue sarcoma. Am J Clin Oncol 2006;29:59–61.

21. Wagner-Bohn A, Paulussen M, Vieira Pinheiro JP et al. Phase II study of gemcitabine in children with solid tumors of mesenchymal and embryonic origin. Anticancer Drugs 2006;17:859–864.

22. Ferraresi V, Ciccarese M, Cercato MC et al. Gemcitabine at fixed dose-rate in patients with advanced soft tissue sarcomas: A mono-institutional phase II study. Cancer Chemother Pharmacol 2008;63:149–155.

23. Merimsky O, Meller I, Kollender Y et al. Gemcitabine in patients with sarcoma of soft tissue or bone resistant to standard chemotherapy. 1999 ASCO Annual Meeting, abst. 2098.

24. Maurel J, Zorrilla M, Puertolas T et al. Phase I trial of weekly gemcitabine at 3-hr infusion in refractory, heavily pretreated advanced solid tumors. Anti-Cancer Drugs 2001;12:713–717.

25. Samuels BL, Barbour L, Schiller D et al. Phase I study of low dose continuous infusion gemcitabine in sarcoma patients. Proc Am Soc Clin Oncol 2001 (abstr 2935).

26. Fox E, Patel S, Wathen JK et al. Phase II study of sequential gemcitabine followed by docetaxel for recurrent Ewing sarcoma, osteosarcoma, or unresectable or locally recurrent chondrosarcoma: Results of Sarcoma Alliance for Research Through Collaboration Study 003. ***The Oncologist*** 2012;17:321.

27. Thall PF, Wathen JK. Practical Bayesian adaptive randomisation in clinical trials. Eur J Cancer 2007;43: 859–866.

28. Korn EL, Freidlin B. Outcome-adaptive randomization: Is it useful? J Clin Oncol 2011;29:771–776.

29. Berry DA. Adaptive clinical trials: the promise and the caution. J Clin Oncol 2011;29:606–609.

See the accompanying article on pages 1213–1220 of this issue.