Cancer Chemother Pharmacol (1994) 34: 361–363

Cancer Chemotherapy and Pharmacology

© Springer-Verlag 1994

Short communication

Response rates – an evolution

Elizabeth Poplin, Laurence Baker

Wayne State University, Internal Medicine, Division of Hematology and Oncology, Detroit, MI 48201, USA

Received: 10 September 1993/Accepted: 9 February 1994

Abstract. The documented frequency of response of cancers to common chemotherapy agents and combinations appears to have decreased over the decades. Multiple reasons exist for this decline including: changes in eligibility and evaluability criteria; changes in the type of patients entered onto trial; and, altered criteria for response and methods for response assessment. The disinclination to publish negative results also permits a bias in a favor of overestimating a drug's efficacy. Circumspection now is advised in assessing data from older trials.

Key words: Response rates – Chemotherapy agents

Deflation, not inflation, is a major concern of chemotherapists. Conscientious observers note that the response rates for chemotherapy agents once considered highly active have dropped. Fluorouracil given by 5-day infusion for colorectal cancer had a reported response rate of 44% in 1975 [1] and 3% in 1990 [2]. The original schedule of 5-fluorouracil (5-FU) given by 5-day bolus was 22% [1] and has more recently been reported as 7% [3]. Mitomycin has been the second-line agent for colorectal cancer because of reported response frequencies ranging up to 33% in small studies conducted from 1968 to 1976 [4–6]. However, in a recent study that utilized mitomycin with or without WR-2721 for the treatment of patients with colorectal cancer, no response was demonstrated among 97 patients, of whom 48 were treated with mitomycin alone and 49 were given the two-agent combination [7]. In the 20 years of their use in the treatment of patients with metastatic soft-tissue sarcomas, the response rate for the combination of bolus doxorubicin and decarbazine has fallen from 41% to 33% to 17% [8–10].

Correspondence to: Elizabeth A. Poplin, M.D., Wayne State University, Division of Hematology and Oncology, Harper Hospital, 3990 John R, Detroit, MI 48201, USA

What has happened to our most traditional chemotherapy regimens? Why are the response rates so much more modest now? There are multiple answers reflecting the markedly different conduct of clinical trials performed now in comparison with those conducted approximately two decades ago, when most of these original data were obtained.

Certainly, some of the change in response rates can be charged to a disinclination toward and delay in publishing negative data. Initial enthusiasm for a new drug may yield a flurry of publications. Data demonstrate that positive results are published more frequently than negative or lackluster results, leading to an imbalance in the perception of a drug's activity [11, 12]. Misplaced initial enthusiasm may not receive a public challenge for many years. These issues have been fully discussed elsewhere [10, 11].

The present discussion focuses on those inherent trial characteristics that relate to eligible patients, study size, patient evaluability, definitions of response, and tools for measuring response. The studies cited above provide examples of all these changes. We anticipate that we can recognize only some of the changes and that others could and will be offered.

As experimental chemotherapists are aware, the population of patients involved in one study is never duplicated in any other study. Some studies differ in size. Others differ in eligibility criteria or patients characteristics. The infusional 5-FU study of 1975 contained 34 patients [1]. The patients had marker lesions consisting of pulmonary nodules, palpable lymph nodes, abdominal masses, or evidence of pelvic or liver involvement with approximately equal frequency. The subsequent study contained 61 patients, of whom 61% had liver as the dominant site of recurrence [2]. Only 16% had pulmonary involvement and 5%, locoregional disease as the dominant site of recurrence. The original doxorubicin-decarbazine trial contained 100 evaluable patients, most of whom had either leiomyosarcoma, osteogenic sarcoma, undifferentiated sarcoma, or fibrosarcoma [8]. Some of the patients had received prior chemotherapy. The most recent sarcoma study, with 118 previously untreated patients, included primarily patients

with leiomyosarcomas and malignant fibrous histiocytomas [10]. Patients with osteogenic sarcomas were specifically excluded.

Furthermore, as surgical techniques as well as combined-modality and adjuvant treatments have become more widely utilized and more successful in the initial treatment of cancers such as sarcomas, head and neck cancers, and gastrointestinal cancers, the population of patients who have been entered into these chemotherapy trials has altered. It is possible that these newly available, aggressive initial therapies cull out those patients with smaller-volume or more sensitive tumors who can be successfully treated. Thus, only those patients with the more aggressive cancers or those with disease resistant to initial combined-modality or adjunctive therapy are entered onto current trials for metastatic disease. These changes in patients' characteristics, inclusion and exclusion criteria, and trial size do not necessarily alter response rates consistently in one direction or another, but they do make each trial unique.

Patients in recent studies have been required to have measurable disease. Prior studies allowed the entry not only of patients with measurable disease but also of those with evaluable disease, hepatomegaly, ascites, jaundice, and bowel obstruction [4]. These latter complications of a patient's cancer were deemed markers of disease that could be followed for progression or response. These latter manifestations of disease would not currently be considered measurable.

Chemotherapy intensity and the correlation with response has previously received much attention and been a subject of excellent and rigorous review [13, 14]. The most recent study of doxorubicin and decarbazine, which had the lowest response rate of three successive trials, employed the same dose of doxorubicin as the other two trials, but the dose of decarbazine delivered was altered from 250 mg/m^2 per day given for 5 days to 750 mg/m^2 given for 1 day only [10]. In this last trial, it is noteworthy that for the subgroup of patients in whom 100% dose therapy was delivered on schedule, the response rate approximated that of the original regimen. However, when doses wee reduced or delayed for reasons other than toxicity (i.e., physician's or patient's request or miscalculation), the response rate dropped significantly. Thus, any interpretation of data from multiple similar trials requires attention to the consistency of dose, schedule, and compliance.

Evaluability criteria have evolved such that fewer patients are excluded from evaluation. For instance, of the 120 patients deemed eligible in the original doxorubicin and decarbazine trial, 19 were deemed inevaluable because of the inability of the patients to complete the first cycle of therapy [8]. Eight patients developed treatment-related toxicity and nine others died during the first cycle. The second sarcoma study stipulated that all patients receiving 5 days of therapy were evaluable [9]. Currently, a study denominator that includes all patients treated is considered standard. This more encompassing denominator includes patients who fail early, patients with chemotoxicity, and patients who refuse or are incapable of receiving further therapy. A response frequency that is based on all patients treated would necessarily be lower than one obtained using more exclusive evaluability criteria, which excludes patients with early progression and those for whom the treatment is too toxic.

The criteria by which a response is judged have become codified, more rigorous, and more technology-based over the years. In a study of mitomycin in 1968, a response required a reduction of $>50\%$ in one or more lesions for two successive measurements [5]. In the study of 1972, for instance, a clinical response included patients in whom there was a "50% decrease in the mass of the measurable lesion . . . opening of a gastrointestinal obstruction and disappearance of jaundice or pleural effusion" [4]. For a partial response, the World Health Organization (WHO) criteria [15] now require a 50% decrease in the sum of the products of the perpendicular diameters of multiple lesions as well as no appearance of new lesions or progression of any existing lesion. The requirements also stipulate that the response must be documented to last for at least 4 weeks. A complete response requires the disappearance of all known disease as determined by two observations undertaken not less than 4 weeks apart. In addition, any abnormal tumor markers must have reverted to normal.

The inaccuracies of repeated reliable physical-examination assessment, necessary for evaluating response, have been noted [16–18]. Studies demonstrate that when measuring simulated or actual masses, physicians may "round off" measurements and underestimate the sizes of masses and nodules. Furthermore, measurements may not be reproducible for a physician measuring the same mass twice or among several physicians measuring the same mass sequentially [16, 17]. It is estimated from these assessments of physicians skills that a false-positive designation of response based on physical examination might occur as often as 10% of the time [16]. In older studies, response determination was often made on the basis of assessment of skin nodules, lymph nodes, clinically palpable tumor masses, liver size, and findings on chest X-ray. In the original 5-FU infusion study [1], for instance, the disease to be assessed consisted of rectal masses, hepatomegaly, intraabdominal tumor masses, skin nodules, lymphadenopathy, pleural effusions, and ascites. Liver scan and chest X-ray were also used.

How much different, then, is the current common practice, in which tumors are measured on computed tomography (CT) scan and WHO criteria are used for defining response. In the recent mitomycin study, CT-scan assessment of liver abnormalities was used for measurement in 72% of the patients; chest X-ray, in 10%; and physical examination in only 2% [7]. No doubt the response rates have been altered by this technological approach to disease assessment. Even so, care must be taken as radiographic measurements may also be subject to errors introduced by differing techniques in sequential radiographic study, including quality of enhancement, imprecise matching of lesions for measurements on serial scans, and imprecise and inconsistent measurements of lesion size.

Tumor markers such as alpha fetoprotein, beta human chorionic gonadotropin (B-HCG), CA-125, and carcinoembryonic antigen (CEA) have become routine markers for germ-cell tumors, ovarian cancer, and gastrointestinal malignancies during the last decade. These markers, if abnormal when the patient's disease is first assessed, must

normalize for documentation of a complete response. These markers provide a very rigorous and sensitive test for assaying response. Similarly, in some studies, documentation of response requires repeat pathologic assessment. For instance, second-look laparotomies have been utilized to document response in ovarian cancer.

In this discussion of response-rate frequency, two caveats must be recognized. In this change to technology-based assessment of tumor response, conventional clinical parameters have been relegated to a second-class status. However, seasoned clinicians surely recognize that patients can receive modest benefit from drugs that cause some decrease in, e.g., hepatomegaly, ascites, and adenopathy, even if the extent of such changes is less than what would be considered necessary to document a partial response. Although it is appropriate to pursue more effective chemotherapy wherever possible, the role of chemotherapy for palliation of symptoms should not be dismissed or forgotten.

If these now rigorously documented responses, both complete (CR) and partial (PR), are to be meaningful or valid, they should translate into a more prolonged response duration. CRs as currently defined are especially difficult to generate, and the survival might be expected to be particularly prolonged. (It is clear that all changes in duration of response that evolve over time are significantly affected by changes that have nothing to do with documentation of the tumor burden but relate instead to factors such as host characteristics and advances in supportive care, among others.) For infusional and bolus 5-FU treatment, the original response duration was 5–6 months [1]. The median duration of response in Kemeny et al.'s 1990 infusion study [2] was 10.4 months. Similar assessment of the three sequential sarcoma studies is difficult. In the first study the remission duration was 4+ months (5+ months for CR); in the second, 31 weeks; and in the third, in which patients were divided into groups according to response category (CR versus PR and including surgical resection), 6.6 months for patients with PRs and 19.6 months for those with CRs. The series of sequential cisplatin, bleomycin, and vinblastine trials for treatment of testicular cancer show a comparable frequency of complete remission over the period ranging from 1974 to 1981 but a larger percentage of patients remaining in prolonged remission with successive trials [19].

Are we then obligated to ignore older data and fastidiously repeat all such studies? This approach would be patient resource-consuming and likely to demonstrate only a modest reduction in the previously noted activity. However, when older data are used as a foundation for some new investigational effort, one need be mindful of the origin of the data and the potential limits of the drugs being used. We should also humbly note that as changes continue to occur in our approach to the practice of oncology, we should anticipate that the future changes in study design and patient assessment may render the current trial data equally outmoded.

References

1. Seifert P, Baker LB, Reed ML, Vaitkevicius VK (1975) Comparison of continuously infused 5-fluorouracil with bolus injection in treatment of patients with colorectal adenocarcinoma. Cancer 36: 123–128
2. Kemeny N, Israel K, Niedzwiecki D, Chapman D, Botet J, Minsky B, et al (1990) Randomized study of continuous infusion fluorouracil versus fluorouracil plus cisplatin in patients with metastatic colorectal cancer. J Clin Oncol 8: 313–318
3. Lokich JJ, Ahlgren JD, Gullo JJ, Philips JA, Fryer JG (1989) A prospective randomized comparison of continuous infusion fluorouracil with a conventional bolus schedule in metastatic colorectal carcinoma: A Mid-Atlantic Oncology Program study. J Clin Oncol 7: 425–432
4. Godfrey T, Wilbur D (1972) Clinical experience with mitomycin in large infrequent doses. Cancer 29: 647–682
5. Moore GE, Bross IDJ, Ausman R, Nadler S, Jones R, Slack N, Rimm A (1968) Effects of mitomycin-C (NSC-26980) in 346 patients with advanced cancers. Cancer Chemother Rep 52: 67–83
6. Baker LH, Izbicki RM, Vaitkevicius VK (1976) Phase II study of porfiromycin vs mitomycin-C utilizing acute intermittent schedules. Med Pediatr Oncol 2: 207–213
7. Poplin EA, LoRusso PDO, Lokich JJ, Gullo JJ, Leming PD, Schulz JJ, Veach SR, McCulloch W, Baker L, Schein P (1994) Randomized clinical trial of mitomycin-C with or without pretreatment with WR-2721 in patients with advanced colorectal cancer. Cancer Chemother Pharmacol 33: 415–419
8. Gottlieb JA, Baker LH, Quagliana JM, Luce JK, Whitecar JP Jr, et al (1972) Chemotherapy of sarcomas with a combination of Adriamycin and dimethyl triazeno imidazole carboxamide. Cancer 30: 1632–1638
9. Baker LH, Frank J, Fine G, Balcerzak SP, Stephens RL, Stuckey WJ, Rivkin S, Saiki J, Ward JJ (1987) Combination chemotherapy using Adriamycin, DTIC, cyclophosphamide, and actinomycin D for advanced soft tissue sarcomas: a randomized comparative trial. A phase III, Southwest Oncology Group study (7613). J Clin Oncol 5: 851–861
10. Zalupski M, Metch B, Balcerzak S, Fletcher WS, Chapman R, Bonnet JD, Weiss GR, Ryan J, Benjamin RS, Baker LH (1991) Phase III comparison of doxorubicin and decarbazine given by bolus versus infusion in patients with soft-tissue sarcomas: a Southwest Oncology Group study. J Natl Cancer Inst 83: 926–932
11. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H Jr (1987) Publication bias and clinical trials. Controlled Clin Trials 8: 343–353
12. Begg CB, Berlin JA (1989) Publication bias and dissemination of clinical research. J Natl Cancer Inst 181: 107–115
13. Hryniuk WM, Figueredo A, Goodyear M (1987) Applications of dose intensity to problems in chemotherapy of breast and colorectal cancer. Semin Oncol 14: 3–11
14. Hryniuk W, Bush H (1984) The importance of dose intensity in chemotherapy of metastatic breast cancer. J Clin Oncol 2: 1281–1287
15. Miller AB, Hoogstraten B, Staquet M, Winkler A (1981) Reporting results of cancer treatment. Cancer 47: 207–214
16. Warr D, McKinney S, Tannock I (1984) Influence of measurement error on assessment of response to anticancer chemotherapy: proposal for new criteria of tumor response. J Clin Oncol 2: 1040–1045
17. Moertel CG, Hanley JA (1976) The effect of measuring error on the results of therapeutic trials in advanced cancer. Cancer 38: 388–394
18. Lavin PT, Flowerdew G (1980) Studies in variation associated with the measurement of solid tumors. Cancer 46: 1286–1290
19. Einhorn CH, Crawford ED, Shipley WU, Loehrer D, Williams SD (1993). Cancer of the testes. In: DeVita (ed) Cancer: principles and practice of oncology, 3rd edn. Lippincott, Philadelphia, p 1138