# Historical Controls for Metastatic Pancreatic Cancer: Benchmarks for Planning and Analyzing Single Arm Phase II Trials

**Philip A. Philip**[1], **Kari Chansky**[2], **Michael LeBlanc**[3], **Lawrence Rubinstein**[4], **Lesley Seymour**[5], **S. Percy Ivy**[6], **Steven R. Alberts**[7], **Paul J. Catalano**[8], and **John Crowley**[9]

[1]Karmanos Cancer Institute, Department of Oncology, Wayne State University, Detroit, MI

[2]Cancer Research And Biostatistics, Seattle, WA

[3]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA

[4]National Cancer Institute, Bethesda, MD

[5]NCIC Clinical Trials Group, Kingston, Ontario Canada

[6]National Cancer Institute, Bethesda, MD

[7]Mayo Clinic, Division of Medical Oncology, Rochester, Minnesota

[8]Dana Farber Cancer Institute, Boston, Massachusetts

[9]Cancer Research and Biostatistics, Seattle, WA

## Abstract

We compiled and analyzed a database of cooperative group trials in advanced pancreatic cancer, to develop historical benchmarks for overall survival (OS) and progression free survival (PFS). Such benchmarks are essential for evaluating new therapies in a single arm setting. The analysis included patients with untreated metastatic pancreatic cancer receiving regimens that included gemcitabine, between 1995 and 2005. Prognostic baseline factors were selected by their significance in Cox regression analysis. Outlier trial arms were identified by comparing individual 6-month OS and PFS rates against the entire group. The dataset selected for the generation of OS and PFS benchmarks was then tested for inter-trial arm variability using a logistic-normal model with the selected baseline prognostic factors as fixed effects and the individual trial arm as a random effect. 1,132 cases from eight trials qualified. Performance status and sex were independently significant for OS, and performance status was prognostic for PFS. Outcomes for one trial (NCCTG-034A) were significantly different from the other trial arms. When this trial was excluded, the remaining trial arms were homogeneous for OS and PFS outcomes after adjusting for performance status and sex. Benchmark values for 6-month OS and PFS are reported along

with a method for using these values in future study design and analysis. The benchmark survival values were generated from a dataset that was homogenous between trials. The benchmarks can be used to enable single-arm phase II trials utilizing a Gemcitabine platform, especially under certain circumstances. Such circumstances might be when a randomized control arm is not practically feasible, an early signal of activity of an experimental agent is being explored such as in expansion cohorts of phase I studies, and in patients who are not candidates for combination cytotoxic therapy.

## Introduction

Phase II clinical trials in cancer have, in recent years, focused increasingly on "targeted" agents that are "cytostatic" rather than "cytotoxic." While most agents that ultimately prove to be useful in the clinic demonstrate at least some disease stability, many authors feel that a traditional treatment response endpoint for phase II trials in solid tumors is less relevant for testing the newer targeted agents (1). Researchers therefore frequently prefer to measure treatment success in terms of overall survival or progression free survival rather than clinical response. For survival and progression free survival endpoints in the phase II setting, one may choose between a single-arm approach which compares trial results with some historical benchmark, or a randomized phase II trial with two or more arms, where the "control" arm provides the benchmark for judging success. The Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee has recommended the randomized approach in the phase II setting, especially when evaluating combinations of agents (2). However, the single-arm approach is deemed appropriate for the evaluation of single agent experimental therapies, and where a well-defined historical control database is available (2,3).

Single-arm designs have the advantage of requiring fewer patients, all of whom receive the experimental treatment. The conduct of trials requires patients, funding, and effort. With a multitude of candidate treatments and limitations on funding and time, an expedited result through a single arm trial is desirable when feasible. However, researchers may have difficulty arriving at an appropriate historical benchmark against which to compare their results (4).

To address the problem of reliable historical benchmarks for single-arm phase II trials, efforts have already been made in specific disease sites, such as stage IV melanoma (3), to amass historical databases and derive historical control data for future trials. The current effort, part of the aforementioned NCI-sponsored task force, has resulted in the compilation of clinical trial data in two specific diseases: advanced pancreatic cancer and advanced non-small cell lung cancer. We report here on the advanced pancreatic cancer database and the benchmarks derived for previously untreated advanced pancreatic cancer. All trials were conducted by cooperative groups in the U.S. from 1995 to 2005. These clinical trial data were compiled and analyzed specifically to provide the appropriate benchmarks for the planning and analysis of future phase II trials in this disease.

Historically, certain trials in advanced pancreatic cancer included locally advanced unresectable disease. More recently, and certainly for the future, trials will select exclusively

for either locally advanced or metastatic disease so that these two patient populations can be studied separately (5). Therefore, in accordance with the primary objective to provide benchmarks for future trials in advanced, metastatic pancreatic cancer, the decision was made to focus initially on cases with metastatic disease and no prior chemotherapy for pancreatic cancer. Cases with locally advanced disease will be considered separately. Additionally, because the use of gemcitabine for advanced pancreatic cancer represented a change in treatment standard for the disease (6), and continues to be a treatment standard, we include for this analysis only those trial arms where gemcitabine is part of the treatment regimen. The recent use of FOLFIRNOX for some patient groups is presented in the discussion.

## Methods

### Data Sources and Selection

Patient-level and trial-level data from eligible cooperative group trial participants were collected from U.S. Cooperative Group phase II and phase III trials in advanced pancreatic cancer, accrued during the time period of January, 1995 through December, 2005. Four cooperative groups submitted data for pancreatic cancer: SWOG, the North Central Cancer Treatment Group (NCCTG), the Eastern Cooperative Oncology Group (ECOG) and the Radiation Therapy Oncology Group (RTOG). The study was approved by the institutional review board for the organization compiling and analyzing the data (Cancer Research And Biostatistics) and by the IRBs serving each participating cooperative group data center. Waiver of consent was approved due to the de-identified and retrospective nature of the data.

Patient-level data elements requested were age at registration, sex, race, Zubrod performance status, location and extent of disease at baseline, prior treatment, laboratory values, and assigned treatment arm. The extent to which the data element requests were fulfilled varied by trial, so that certain baseline factors were missing from various subsets of the full dataset. Trial-level variables included activation and closure dates, the treatment setting (first line versus second or more), details on the treatment regimen, type of trial (phase II versus III), objective tumor response and disease progression criteria in use, and type of disease evaluation required at baseline. Patient-level outcome data included best response to treatment, time to progression, and time to death or last contact. Trials activating prior to the 2000-2004 time frame utilized bi-dimensional response and progression criteria as described below. Those activated 2000 or later utilized uni-dimensional RECIST (7).

### Statistical Methods

Overall survival was calculated as the time from registration until death or last contact, with censoring at last contact if the patient was still living. Progression free survival was calculated as the time from registration until documented progression of disease (according to individual trial criteria), or death, whichever occurred first. Cases alive without progression were censored at the date of last contact.

The selection of baseline patient-level factors to carry forward as part of the determination of overall and progression free survival benchmarks was achieved by the combined information from statistical modeling and practical considerations. Practical considerations included the availability of a given data element in a sufficient proportion of the submitted data, and the practicality of measuring or predicting the distribution of a given factor in a real trial population. The selection of independently prognostic baseline covariates was achieved via Cox regression (8) with stepwise variable selection, applied separately to overall survival (OS) and progression free survival (PFS). Univariately significant patient-level factors with sufficient representation in the database were subjected to multiple factor Cox regression with stepwise selection to determine independent prognostic significance. Factors that were not univariately significant (but sufficiently represented) were initially included as well, in order to explore for the possibility of confounding. Factors that were insufficiently represented in the data were excluded from consideration. As a rule, a factor that was represented in less than 80% of the cases available for analysis, or was missing completely from an individual contributing group in the final analysis set, could not be considered. Survival curves and estimates of medians and rates were estimated by the Kaplan-Meier method (9).

Between-trial variability in the six-month overall survival rate and the 6 month progression free survival rate was assessed in two ways: 1) the OS and PFS rate for each trial arm was compared to the overall rate using a Bonferroni correction; 2) once prognostic baseline factors were selected, a logistic normal model with individual trial arm as a random effect was used to estimate between-trial variability not accounted for by differing distributions of the baseline factors. These methods for assessing between-trial variability were also described in the context of historical melanoma trials by Korn et al. (3). This model was applied to the entire set of trials, as well as to the reduced set where outliers identified in the first method were eliminated. Overall and progression free survival benchmarks were calculated for prognostic groups according to the selected baseline factors, eliminating any trial that was found to be an outlier in step 1. All analyses were performed using SAS version 9.2.

## Results

### Characteristics of the Study Population

A complete listing of the trials involving 2,341 cases that were contributed by four cooperative groups is given in (Table 1). After screening to include only those cases with previously untreated metastatic disease enrolled to trial arms that incorporated gemcitabine as part of the regimen, there were 1,132 cases available for this analysis. The included cases were from eight trials (ten trial arms) with activation dates ranging from 1996 to 2004, as characterized in Table 2. The three ECOG trials that were activated prior to the year 2000 utilized ECOG solid tumor response criteria (10). With respect to determining progression, the ECOG criteria are comparable to standard WHO bi-dimensional progression criteria (11), in that a 25% increase in the product of perpendicular diameters in any single lesion (or the appearance of a new lesion) qualified for disease progression. The NCCTG trial activated prior to the year 2000 used comparable bi-dimensional criteria. All other trials

used RECIST, which for progression requires a 20% increase in the sum of uni-dimensional diameters over the smallest sum observed during treatment, or the observation of new lesions. Assessment intervals ranged from 6 weeks to 8 weeks (or two cycles of gemcitabine). Reported outcomes for each included trial arm are given in Table 3.

Of the 1,132 cases, 1,123 had died and nine were alive at last contact. The minimum, median, and maximum survival followup for the 9 living patients were 2 months, 43 months, and 74 months, respectively. All but 3 of the 1,132 patients had progressive disease at the time of death or last contact, or had died due to their cancer. Published results of individual trials are shown in Table 3.

Baseline characteristics for the 1,132 cases are shown in Table 4. Forty-six percent of patients were female, and a majority of patients (87%) had a performance status (translated to Zubrod scale) of 0 or 1. Ninety percent of patients were Caucasian, 7% black, <1% Asian, and the remaining 2% Native American, Pacific Islander, or not reported. Seventy-seven percent of cases originated from randomized phase III trials. Thirty-eight percent of cases were from trials that were activated from the period of 1995-2000, 62% from 2000-2005.

## Prognostic Baseline Factors

Univariate survival statistics for the considered baseline factors with respect to overall and progression free survival are shown in Table 4. Baseline factors considered were: sex, race (white versus African American versus Asian versus other), performance status, bilirubin > normal, and serum SGOT/SGPT > normal. Location of metastatic disease was not considered due to insufficient data availability. Certain laboratory values such as albumin and platelet counts were also not consistently available across the submitted databases.

Overall survival findings were as follows: Race was not univariately significant as no single race was significantly different from the others, possibly because the distribution was overwhelmingly white. The race factor was explored in the multivariate setting initially, to assess for the possibility of effect modification with other factors. This factor was then eliminated from consideration. SGOT/SGPT was significant univariately (HR 1.35, P<.001 for abnormal SGOT/SGPT) but was not entered into multivariate analyses because of insufficient numbers with available data (available for only 839 cases.) Serum bilirubin level was significantly prognostic for survival (HR = 1.52 for bilirubin above normal, p<.001), but was missing in 540 (nearly half) of cases. Imputation for patient data with missing labs was considered but lacked good surrogates for an imputation model. The remaining factors (sex and performance status) were entered into a Cox regression analysis with stepwise elimination, with the final model including both factors as independently associated with survival.

For progression free survival, race, sex, and bilirubin were not significantly prognostic. SGOT/SGPT was significant (HR 1.34, P<.001), but again, the factor was not considered further due to lack of sufficient data. Performance status was independently prognostic (Table 4).

### Variability Between Trial Arms

The overall six-month survival rate for the combined 10 trial arms was 48% (95% C.I. 44% - 50%). When comparing the individual six month overall survival rates of each of the 10 trial arms to the overall rate of 48% with correction for multiple comparisons, one trial arm – the single arm phase II trial NCCTG-034A (gemcitabine + bevacizumab + oxaliplatin), with a 6-month survival rate of 67% for 70 patients – differed substantially from the group. A logistic normal model for the binary six-month overall survival outcome was applied with performance status and male sex included as fixed effects, and a random effect with an assumed normal distribution to represent the residual variance component not explained by the two baseline covariates. The P-value for the variance component (by likelihood ratio test) was 0.032, indicating significant inter-trial-arm variability. The same model after excluding the outlier trial NCCTG-034A yielded a non-significant p-value of 0.14 for the inter-trial variance component. Thus with respect to variability between trial arms in six-month overall survival rates, the removal of NCCTG-034A serves to achieve the homogeneity for reliable benchmarks.

For the six-month progression free survival rates, a comparison of each arm to the overall 6-month PFS rate of 24% (95% C.I. 22% - 27%) identified the same outlier, NCCTG-034A, with a 6 month PFS rate of 44%. There was a significant variance component in the logistic-normal model even after adjusting for the covariate of performance status (p=.02), representing considerable between trial-arm variability. When the outlier trial was removed from the dataset, the between trial-arm variance component was no longer significant (p is close to 1 when adjusted for performance status). Based on these findings, a suitable benchmark for progression-free survival might be derived using the reduced dataset, which, again, excludes the outlying trial arm from NCCTG-034A.

### Survival and Progression Free Survival

The median overall survival for all trials combined, henceforth excluding NCCTG-034A, was 5.7 months (95% C.I. 5.3 – 6.0 months), and the median progression free survival was 2.9 months (95% C.I. 2.6 – 3.4 months). There were no differences in survival or progression free survival when comparing the data from Phase II vs. Phase III trials. (P=.70 for OS and P=.67 by log-rank test for PFS; Figure 1A and B). Likewise, there were no significant differences in OS or PFS when comparing the 1995-2000 activation period against later 2000-2004 activations (Figure 1C and D). Benchmarks for OS and PFS according to the chosen factors (sex and performance status) are given in Tables 5 and 6, respectively. The subgroups with good performance status (0 or 1 on the Zubrod scale) had the better prognoses for overall and progression free survival, and females had a slight overall survival advantage over males. Although sex was not a statistically significant factor in the Cox regression analysis for PFS, we chose to include this factor in the derivation of benchmarks for PFS to maintain consistency with the overall survival benchmarks. Similarly, predicted survival rates for cases classified by sex and performance status are provided for and utilized in the study planning and analysis examples given in Examples 1 and 2. Figure 2A and B show overall and progression free survival curves for the same benchmark categories.

**Application of the benchmark algorithm for future phase II trial designs**

The variation in PFS or OS based on the prognostic factors indicated above (sex and performance status) can be combined to arrive at benchmarks (or null hypothesis values) for future clinical trials. Estimates from Tables 5 and 6 above can be used to predict survival for patients registered to a new study, depending on the specific proportions with respect to the prognostic factors. Example 1 shows the procedure for designing a study with the 6 month OS rate as primary endpoint. One can estimate (or guess) the potential frequency of patient groups to come up with an appropriate benchmark for the expected study sample. The guesses will not dramatically impact the appropriate sample size. The expected fraction of patients in each of the (sex, performance status) groupings are multiplied by the 6 month OS estimates from Table 5 (following the algorithm of Korn et al. [3]) in order to generate a null predicted 6-month (or 12-month) survival rate:

**Example 1**: Designing a single-arm phase II trial with 6 month OS rate as primary endpoint.

1. Estimate expected frequencies of patient groups to be accrued, to determine $\pi_{P0}$, the benchmark null hypothesis:

   Example:

   Female, PS 0 (5%); Female, PS 1 (30%); Female, PS 2 (10%)

   Male, PS 0 (5%); Male, PS 1 (40%); Male, PS 2 (10%)

2. Use predicted rates in table 5 to calculate the predicted 6-month OS rate for this sample.

   Using the above frequencies:

   $\pi_{P0}$=.05$x$.64+.30$x$.45+.10$x$.26+.05$x$.57+.40$x$.47+.10$x$.27 = .44 or 44%.

3. Specify the alternative $\pi_{PA}$.

   Example: $\pi_{PA=}$ $\pi_{P0}$+.15.

4. Use usual binomial sample size calculators with assumption of complete followup at 6 months.

Example 2 shows the procedure for the analysis of a completed trial given the actual fraction of patients in each grouping. The calculated predicted OS rate is compared to the observed 6 month survival estimated from the new Phase II trial. A new treatment could be declared worthy of additional study if the OS rate $\pi_P$ can be rejected at some type-1 error level (for instance, p<.10):

**Example 2**: Analyzing a completed single-arm phase II trial with 6 month OS rate as primary endpoint:

1. Use the actual frequencies of patient groups to determine $\pi_P$, the average of predicted outcomes specific to the trial.

   Example:

   Female, PS 0 (5%); Female, PS 1 (30%); Female, PS 2 (10%)

Male, PS 0 (5%); Male, PS 1 (40%); Male, PS 2 (10%)

2.  Use predicted rates in table 5 to calculate the historical null rate prediction for OS.

    Using the example expected frequencies:

    $\pi_P$=.05x.64+.30x.45+.10x.26+.05x.57+.40x.47+.10x.27 = .44 or 44%.

3.  Compare the predicted rate to the observed 6 months OS rate in the completed trial.

    Declare the treatment worthy if the predicted rate (0.44) can be rejected at the desired type 1 error level (e.g., p<.10).

These examples represent a trial patient population with a slightly worse performance status distribution than that seen in the historical data set and the 6 month predicted OS is adjusted to that patient mixture.

## Discussion

With the advent of newer targeted therapies there is potential for improvement over existing cytotoxic therapies in the treatment of pancreatic cancer (12, 13). The Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee recommended the implementation of a randomized approach for the phase II setting in general, but acknowledged the existence of situations where a single-arm approach is allowable (2). Specifically, if there exists a reliable historical database, the single-arm approach may be a way to reduce the number of patients required to complete the study (4).

The availability of appropriate data to serve as a historical control provides a mechanism to help design and assess the addition of these new agents initially in a phase II setting when a single arm trial is desired, or multiple treatment arms all consisting of experimental therapies. Accrual may even proceed more quickly if the aspect of randomization diminishes the desire to participate for some patients where existing therapies have marginal benefit. In fact a large national survey of cancer patients shows that the most frequent barrier to trial accrual, on the part of the patients, was their concern about randomized treatment (14).

With a modern, relevant historical control database for untreated, metastatic pancreatic cancer, the next step would be to create online tools, based on the historical benchmarks, to be used for study planning and analysis of single-arm or multiple-arm "pick the winner" phase II trials.

There was complete overlap in terms of survival between the phase II and phase III trial setting in the benchmark population, which supports the use of phase III trial data to supplement the availability of phase II in the establishment of benchmarks for phase II trials. Progression criteria varied in the study populations with either some version of bi-dimensional response and progression criteria or RECIST in use on any given trial. RECIST is generally thought to be less sensitive to disease progression because the volume required to call progressive disease is larger and patients satisfying the criteria for progression under WHO (with a 25% increase in the bi-dimensional product in any single lesion) may not satisfy the criteria for progression under RECIST. One might therefore expect shorter progression free survival times in the trials activating prior to 2000. Assessment intervals

varied as well, with the shortest protocol-specified interval being 6 weeks, and the longest 8 weeks that corresponded to two cycles of gemcitabine. For this reason, overall survival remains as the most reliable endpoint in pancreatic cancer trials in general, where progression free survival times are frequently short enough to be biased by differing assessment intervals (5). Additionally, with very short survival times, overall survival is a reasonable endpoint for advanced pancreatic cancer even in the phase II setting where results are expected in a shorter time frame.

Despite the potential effects on progression times based on progression criteria, in this dataset there was no significant difference in PFS between older trials, activated during the era of bi-dimensional progression criteria, versus newer trials using RECIST. This would suggest that the PFS endpoint is comparable even when criteria differ. However, this topic deserves further exploration, including a study of the comparability of solid tumor response endpoints, and this database will enable that exploration as well.

Estimates of the overall survival and progression free survival rates across all of the trials would be misleading if there were significant between-trial variance that could not be explained by differences in the population with respect to the chosen baseline factors (15). In this case, the elimination of one outlying trial resulted in a population without a significant between-trial variance component.

There are several compelling reasons to further this effort, once the necessary tools are in place, by continuing to build a database as trials are completed. If the database can be constantly updated with new trials, it will remain an invaluable tool for planning and analysis. Changes in treatment standards and improvements in first, second-line therapy will ultimately result in benchmarks in need of continuous updates. Although there was no difference in survival over the past ten years in the dataset used for the primary analysis, this was not so for the entire database. For the data initially contributed to this project, the decision was made to include only those trials that administered gemcitabine as part of the regimen, because this represented a change in treatment standard and an accompanying advance in the survival prognosis (6, 12).

The "outlier" trial NCCTG-034A, excluded from the benchmark calculations because its results differed markedly from the rest of the database in the positive direction. The recent development of the non-gemcitabine based cytotoxic combination regimen of oxaliplatin, irinotecan, 5-fluorourcil and leucovorin (FOLFIRINOX) (16) demonstrated improved survival of patients with metastatic disease over gemcitabine. However, the applicability of the FOLFIRINOX regimen, usually in a modified form, is limited to 20-25% of patients with metastatic disease who have a favorable performance status (0-1) and adequate liver function and are generally younger. For now, gemcitabine-based regimens remain the standard of care for the majority of patients with advanced pancreatic cancer and the database reported upon here remains relevant for the future. A recent phase III study demonstrated the superiority of nab-paclitaxel/gemcitabine combination when compared to gemcitabine that would support the use of gemcitabine-based therapies in advanced pancreatic cancer in the foreseeable future (17). Nevertheless, future additions to a database of historical controls would be a key to maintaining its relevance. Further work on

evaluating the performance of the model on recent completed and future Phase II trial results would also be interesting. Expanding these analyses to other more comprehensive datasets such as the Aide et Recherche en Cancerelogie Digestive (ARCAD) pancreatic database would also be a very important consideration.

The discovery of factors that are prognostic for survival in pancreatic cancer was not the purpose of this study. However, there are other patient-level baseline factors in addition to performance status and sex that are undoubtedly prognostic for survival and progression free survival, and it may be practical for some of these factors to be used in the refinement of historical benchmarks. For example, although LDH was requested it was not available from the contributing groups. Based upon initial review of availability, factors such as percentage weight loss, serum albumin concentration, serum CA19-9 level, and tumor grade were not requested though there is some evidence reported that they might be prognostic. Location of the primary tumor and metastatic sites could also be considered as factors in the future. With a more extensive database those factors could be found and utilized for the further development of historical benchmarks that tailors more closely to a specific phase II trial's patient population. This would increase the advantage to this covariate-adjusted approach, as compared to the simpler method of just comparing overall outcomes with historical benchmarks. Furthermore, other groups represented in the database in smaller proportions, such as those with locally advanced disease treated with or without radiotherapy in conjunction with chemotherapy, are worthy of investigation in order to develop benchmarks applicable to those populations.

Continued development of a historical control database would lend itself well to the development of web-based study planning and analysis tools. These freely accessible tools would enable researchers to design an appropriately powered single- or multiple-experimental arm studies in much the same way as currently available tools provide. The enhancement would be in their ability to specify the expected patient population in terms of the baseline factors. They will also enable the researcher to perform analyses at the conclusion of the trial, which account for the actual makeup of the trial population with respect to the important baseline factors, allowing for a comparison against an appropriately matched historical control. It is worth noting that these target values could be used to help develop trial designs other than single arm studies. For example, they could be used to provide baseline survival calculations to aid in the design for randomized selection or screening studies that do not incorporate a control arm. If the database continues to grow, these tools can be continuously refined. As these tools are developed and used, it is hoped that the experience gained can be successfully applied to other disease settings in phase II cancer clinical trials.

## Acknowledgments

## References

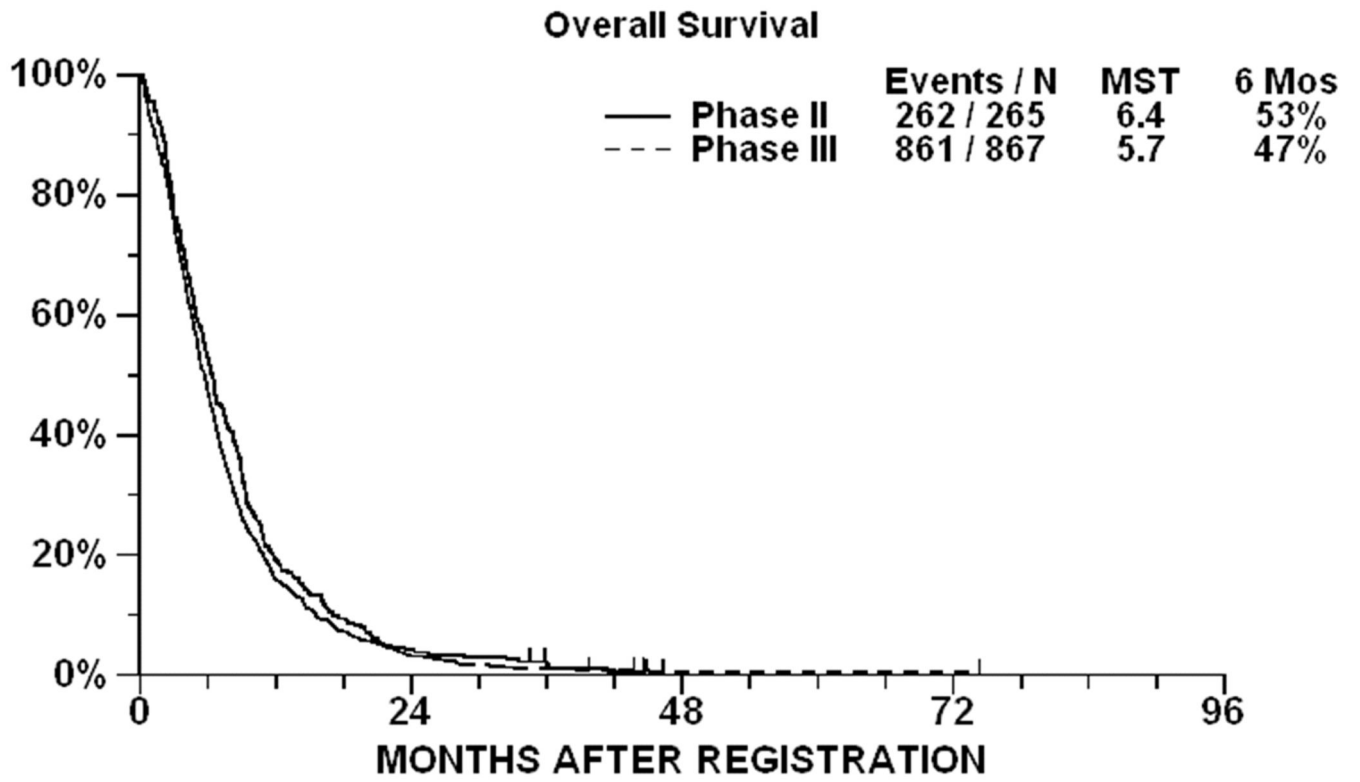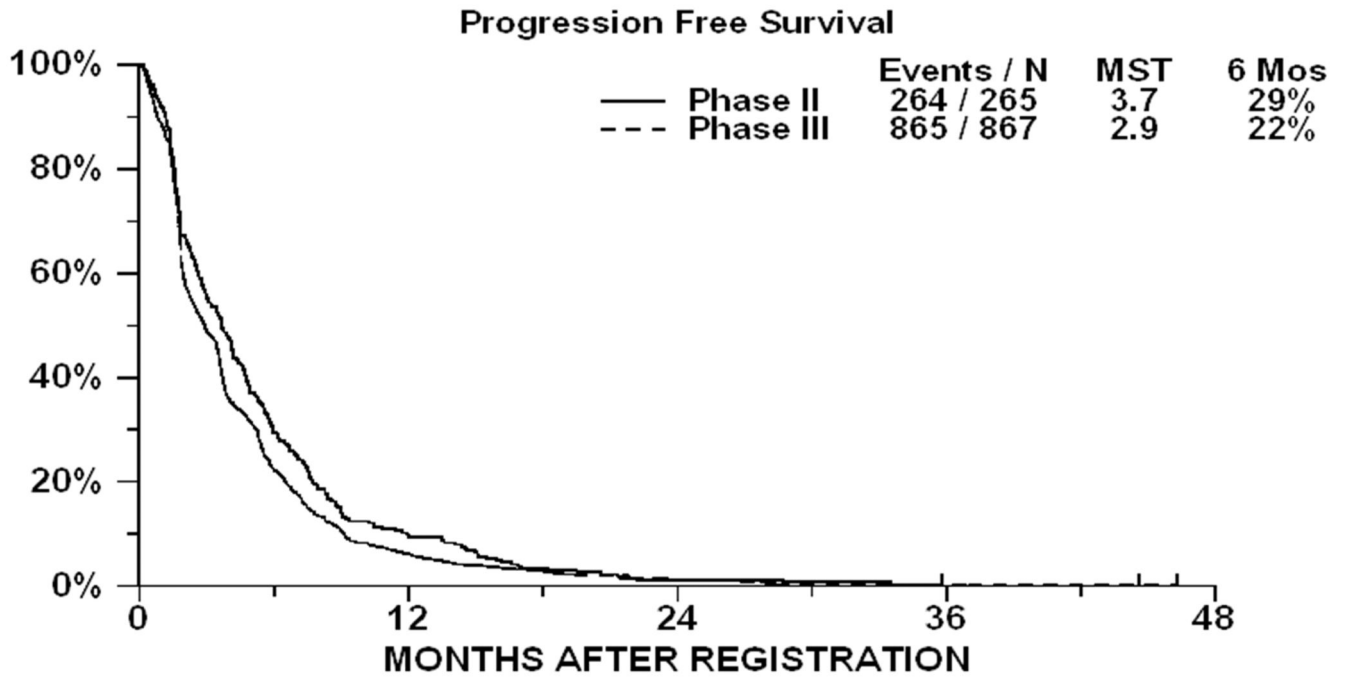1. Dhani N, Tu D, Sargent DJ, Seymour L, Moore MJ. Alternate endpoints for screening phase II studies. Clin Canc Res. 2009; 15:1873–82.

2. Seymour L, Ivy SP, Sargent D, Spriggs D, Baker L, Rubinstein L, et al. The design of phase II clinical trials testing cancer therapeutics: Consensus recommendations from the Clinical Trial Design Task Force of the National Cancer Institute Investigational Drug Steering Committee. Clin Canc Res. 2010; 16:1764–9.

3. Korn EL, Liu PY, Lee SJ, Chapman JA, Niedzwiecki D, Suman VJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. J Clin Oncol. 2008; 26(4):527–34. [PubMed: 18235113]

4. Rubinstein L, Crowley J, Ivy P, LeBlanc M, Sargent D. Randomized phase II designs. Clin Canc Res. 2009; 15:1883–90.

5. Philip P, Mooney M, Jaffe D, Eckhardt G, Moore M, Meropol N, et al. Consensus report of the National Cancer Institute Clinical Trials Planning Meeting on Pancreas Cancer Treatment. J Clin Oncol. 2009; 27:5660–9. [PubMed: 19858397]

6. Burris HA, Moore MJ, Andersen J, Green MR, Rothenberg ML, Modiano MR, et al. Improvements in survival and clinical benefit with gemcitabine as first-line therapy for patients with advanced pancreas cancer: A randomized trial. J Clin Oncol. 1997; 15:2403–13. [PubMed: 9196156]

7. Therasse P, Arbuck S, Eisenhauer E, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000; 92:205–16. [PubMed: 10655437]

8. Cox D. Regression Models and Life-Tables. J R Stat Soc Series B Stat Methodol. 1972; 34:187–220.

9. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53:457–81.

10. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American J Clin Oncol. 1982; 5:649–55.

11. WHO handbook for reporting results of cancer treatment. World Health Organization Offset Publication; Geneva, Switzerland: 1979.

12. Burris H, Rocha-Lima C. New therapeutic directions for advanced pancreatic cancer: targeting the epidermal growth factor and vascular endothelial growth factor pathways. Oncologist. 2008; 13:289–98. [PubMed: 18378539]

13. Heinemann V, Boeck S, Hinke A, Labianca R, Louvet C. Meta-analysis of randomized trials: evaluation of benefit from gemcitabine-based combination chemotherapy applied in advanced pancreatic cancer. BMC Cancer. 2008; 8:82. [PubMed: 18373843]

14. Unger JM, Hershman DL, Albain KS, Moinpour CM, Petersen JA, Burg K, et al. Patient income level and cancer clinical trial participation. J Clin Oncol. 2013; 31:536–42. [PubMed: 23295802]

15. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2008; 172:137–59.

16. Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouarn Y, et al. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. N Engl J Med. 2011; 364:1817–25. [PubMed: 21561347]

17. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante JR, Moore MJ, et al. Randomized Phase III Study of Weekly *nab*-Paclitaxel plus Gemcitabine vs Gemcitabine Alone in Patients with Metastatic Adenocarcinoma of the Pancreas, in 2013 Gastrointestinal Cancer Symposium. J Clin Oncol. 2012; 30(suppl 34)

18. Shepard RC, Levy DE, Berlin JD, Stuart K, Harris JE, Aviles V, et al. Phase II study of gemcitabine in combination with docetaxel in patients with advanced pancreatic carcinoma (E1298). Oncology. 2004; 66:303–9. [PubMed: 15218298]

19. Berlin JD, Catalano P, Thomas JP, Kugler JW, Haller DG, Benson A. Bowen III. Phase III study of gemcitabine in combination with fluorouracil versus gemcitabine alone in patients with advanced pancreatic carcinoma: Easter Cooperative Oncology Group trial E2297. J Clin Oncol. 2002; 20:3270–5. [PubMed: 12149301]

20. Berlin JD, Adak S, Vaughn DJ, Flinker D, Blaszkowsky L, Harris JE, et al. A phase II study of gemcitabine and 5-fluorouracil in metastatic pancreatic cancer: An Eastern Cooperative Oncology Group study (E3296). Oncology. 2000; 58:215–8. [PubMed: 10765123]

21. Alberts SR, Schroeder M, Erlichman C, Steen PD, Foster NR, Moore DF Jr. et al. Gemcitabine and ISIS-2503 for patients with locally advanced or metastatic pancreatic adenocarcinoma: A North Central Cancer Treatment Group Phase II Trial. J Clin Oncol. 2004; 22:4944–50. [PubMed: 15611509]

22. Alberts S, Foster N, Morton R, Kugler J, Schaefer P, Wiesenfeld M, et al. PS-341 and gemcitabine in patients with metastatic pancreatic adenocarcinoma: a North Central Cancer Treatment Group (NCCTG) randomized phase II study. Ann Oncol. 2005; 16:1654–61. [PubMed: 16085692]

23. Kim GP, Oberg A, Kabat B, Sing A, Hedrick E, Campbell S, et al. NCCTG phase II trial of bevacizumab, gemcitabine, oxaliplatin in patients with metastatic pancreatic adenocarcinoma. J Clin Oncol, 2006 ASCO Meeting Proceedings Part I. 2006; 24(18s)

24. Alberts S, Townley P, Goldberg R, Cha SS, Sargent DJ, Moore DF, et al. Gemcitabine and oxaliplatin for metastatic pancreatic adenocarcinoma: a North Central Cancer Treatment Group phase II study. Ann Oncol. 2003; 14:3605–10.

25. Philip PA, Benedetti J, Corless C, Wong R, O'Reilly EM, Flynn PJ, et al. Phase III study comparing gemcitabine plus cetuximab versus gemcitabine in patients with advanced pancreatic adenocarcinoma: Southwest Oncology Group-directed intergroup trial S0205. J Clin Oncol. 2010; 28:3505–10.

A

## Overall Survival

| | Events / N | MST | 6 Mos |
|---|---|---|---|
| —— Phase II | 262 / 265 | 6.4 | 53% |
| - - - Phase III | 861 / 867 | 5.7 | 47% |



MONTHS AFTER REGISTRATION

B

## Progression Free Survival

| | Events / N | MST | 6 Mos |
|---|---|---|---|
| Phase II | 264 / 265 | 3.7 | 29% |
| Phase III | 865 / 867 | 2.9 | 22% |

MONTHS AFTER REGISTRATION

C

## Overall Survival

|  | | Events / N | MST | 6 Mos |
|---|---|---|---|---|
| —— | 1995-1999 | 422 / 423 | 5.7 | 48% |
| - - - | 2000-2004 | 633 / 639 | 5.6 | 47% |
| — — | 2005 | 68 / 70 | 8.7 | 67% |

MONTHS AFTER REGISTRATION

D

## Progression Free Survival

| | | Events / N | MST | 6 Mos |
|---|---|---|---|---|
| —— | 1995-1999 | 423 / 423 | 2.8 | 24% |
| - - - | 2000-2004 | 637 / 639 | 3.0 | 21% |
| – – | 2005 | 69 / 70 | 5.7 | 44% |

**Figure 1.**
A. Overall survival according to type of clinical trial (Phase II vs. Phase III). Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,062. (N034A excluded.) B. Progression free survival according to type of clinical trial (Phase II vs. Phase III). Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,062. (N034A excluded.) C. Overall survival according to time period of trial activation. Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,062. (N034A excluded.) D. Progression free survival according to time period of trial activation. Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,062. (N034A excluded.)

A

## Overall Survival



|  | Events / N | MST | 6 Mos |
|---|---|---|---|
| Fem.PS 0 | 140 / 144 | 7.9 | 65% |
| Fem.PS 1 | 295 / 298 | 5.6 | 46% |
| Fem.PS 2 | 69 / 69 | 3.8 | 26% |
| Male PS | 603 / 605 | 5.8 | 48% |

MONTHS AFTER REGISTRATION

B

## Progression Free Survival

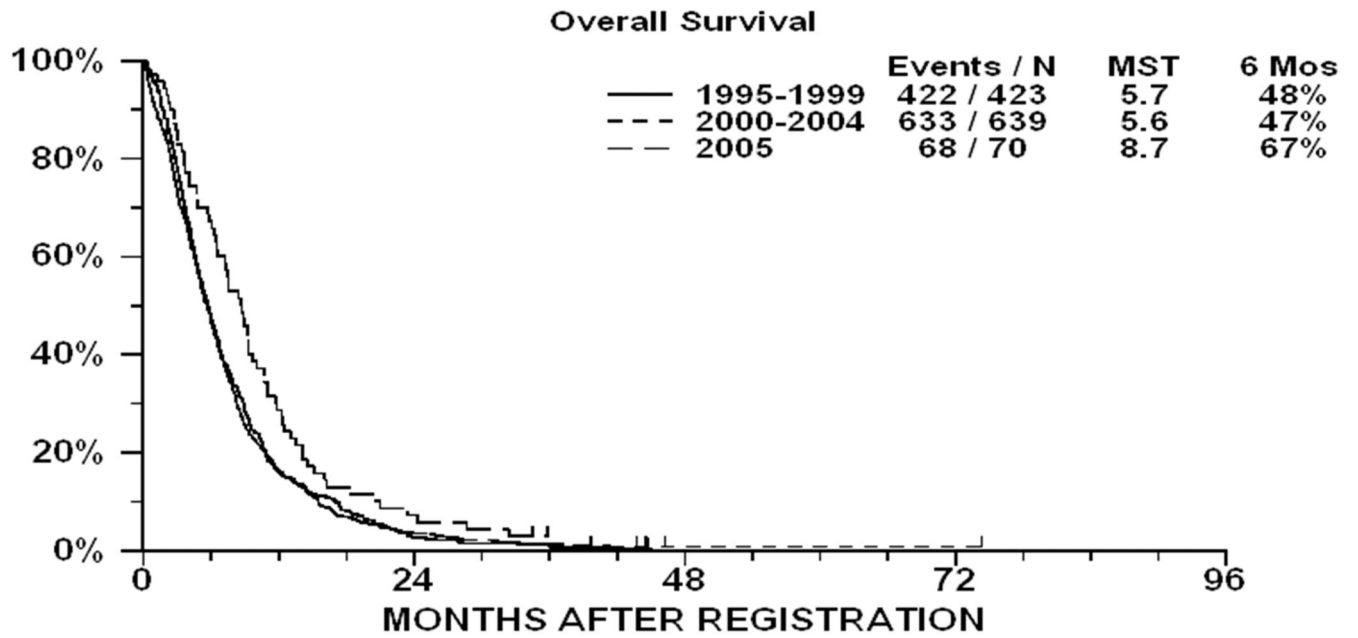| | Events / N | MST | 6 Mos |
|---|---|---|---|
| Fem.PS 0 | 141 / 144 | 3.8 | 28% |
| Fem.PS 1 | 298 / 298 | 3.5 | 26% |
| Fem.PS 2 | 69 / 69 | 2.1 | 13% |
| Male PS | 605 / 605 | 2.8 | 23% |
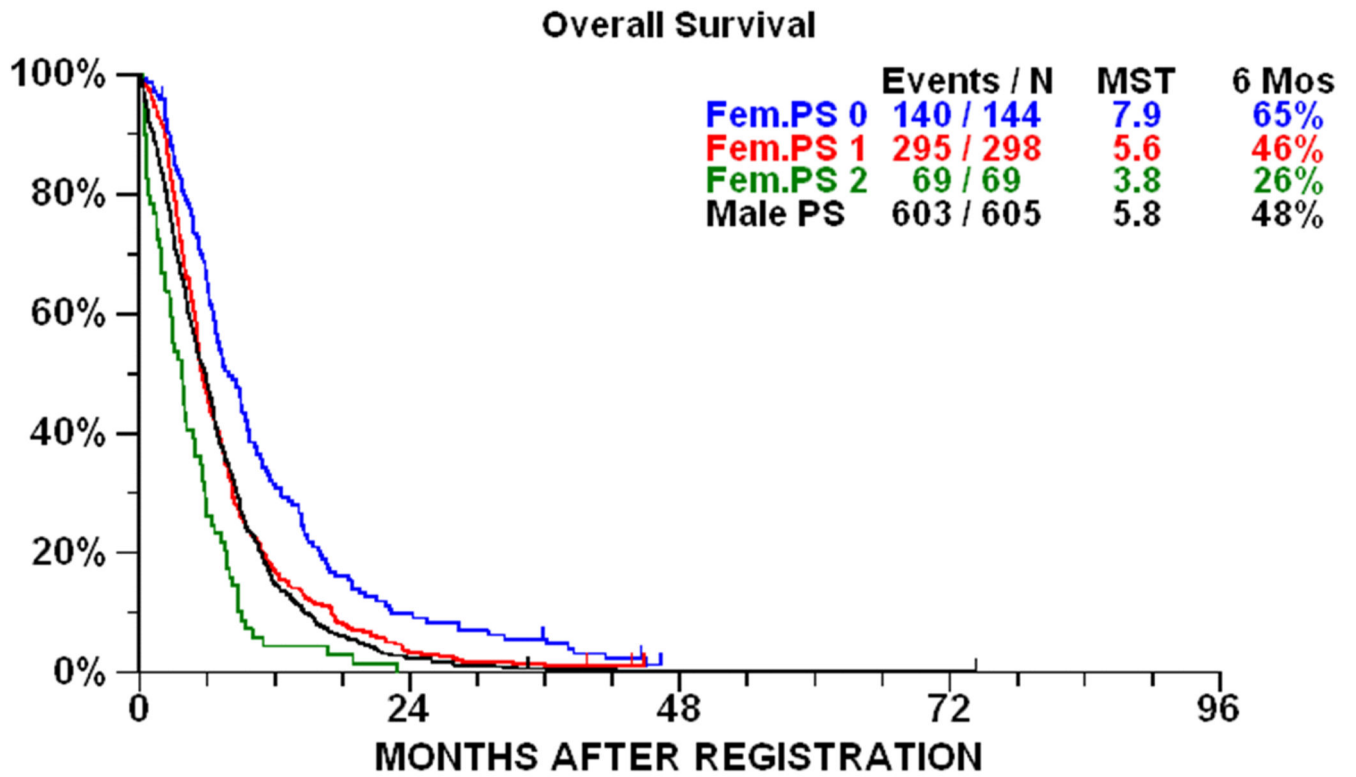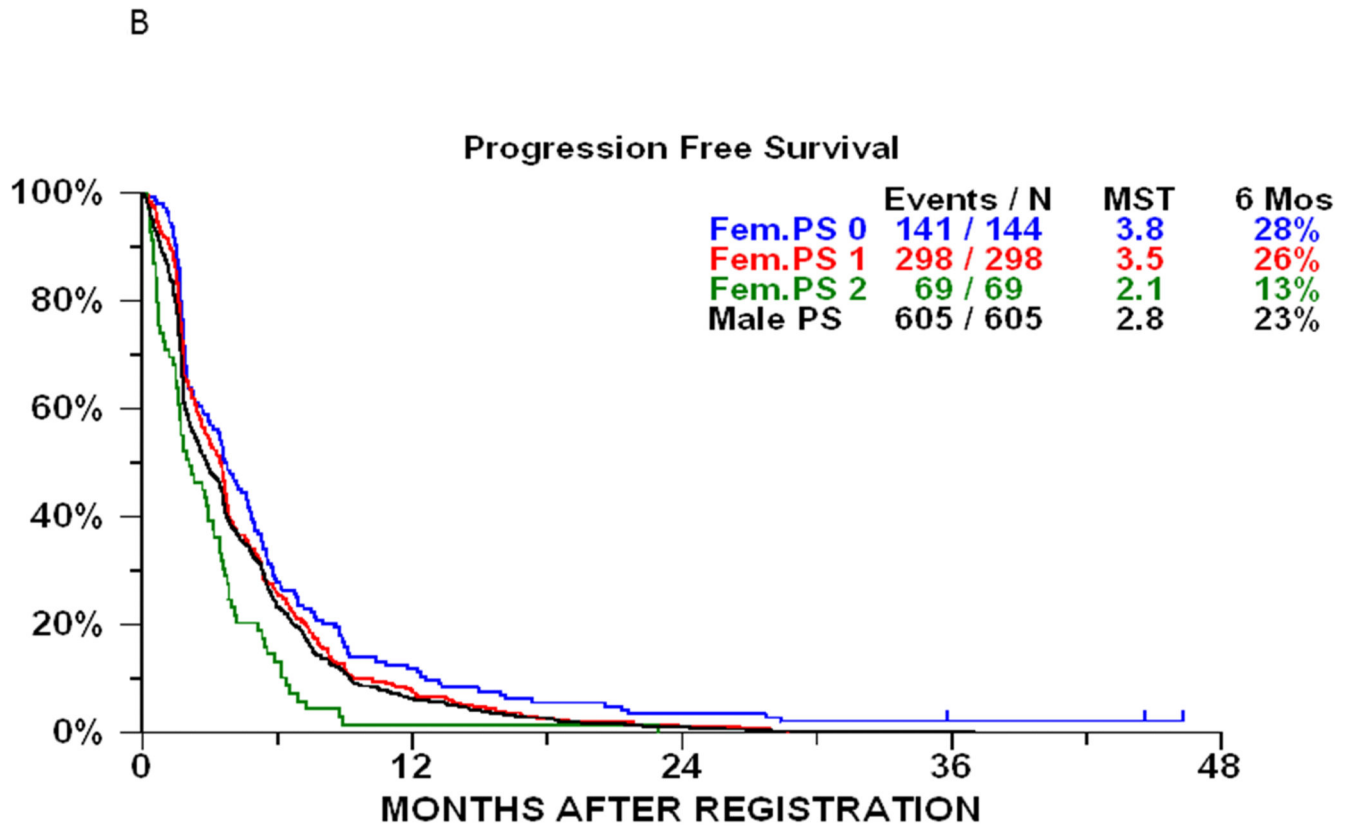
**Figure 2.**
A. Overall survival according to sex and performance status. Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,046 cases with complete data for sex and performance status. B. Progression free survival according to performance status and sex. Cooperative group trials in advanced pancreatic cancer with regimens containing Gemcitabine. N=1,046 cases with complete data for performance status, NCCTG-034A excluded. Abbreviation: PS, performance status.

**Table 1**

Complete Listing of Data Contributed by Group and Protocol:

| Protocol | Type of Trial | Protocol Tx | Setting | N |
|---|---|---|---|---|
| ECOG-1202 | Phase II | Chemotherapy | First Line | 10 |
| ECOG-1298 | Phase II | Chemotherapy | First Line | 32 |
| ECOG-2297 | Phase III | Chemotherapy | First Line | 321 |
| ECOG-3292 | Phase II | Chemo/Immunoth. | First Line | 24 |
| ECOG-3296 | Phase II | Chemotherapy | First Line | 36 |
| NCCTG-0043 | Phase II | Chemo/Biologic | First Line | 48 |
| NCCTG-014C | Randomized Phase II | Biologic | First Line | 42 |
| NCCTG-014C | Randomized Phase II | Chemo/Biologic | First Line | 39 |
| NCCTG-034A | Phase II | Chemo/Biologic | First Line | 79 |
| NCCTG-894352 | Phase III | Chemotherapy | Either | 94 |
| NCCTG-924352 | Phase II | Chemotherapy | Either | 46 |
| NCCTG-964351 | Phase II | Chemotherapy | First Line | 13 |
| NCCTG-984351 | Phase II | Chemotherapy | Either | 58 |
| NCCTG-9942 | Phase II | Concurrent Chemo/RT | First Line | 46 |
| RTOG-0020 | Randomized Phase II | Concurrent Chemo/RT | First Line | 91 |
| RTOG-0411 | Phase II | Chemo/biologic/RT | First Line | 82 |
| RTOG-9102 | Phase III | Concurrent Chemo/RT | First Line | 27 |
| RTOG-9209 | Phase II | Concurrent Chemo/RT | First Line | 50 |
| RTOG-9812 | Phase II | Concurrent Chemo/RT | Either | 109 |
| SWOG-0107 | Phase II | Chemotherapy | First Line | 60 |
| SWOG-0205 | Phase III | Chemotherapy | First Line | 347 |
| SWOG-0205 | Phase III | Chemo/Biologic | First Line | 348 |
| SWOG-8916 | Phase II | Chemotherapy | First Line | 25 |
| SWOG-8933 | Phase II | Chemotherapy | First Line | 35 |
| SWOG-9100 | Phase II | Chemo/Potentiator | First Line | 26 |
| SWOG-9135 | Phase II | Chemotherapy | First Line | 39 |
| SWOG-9413 | Phase II | Chemo/Immunoth. | First Line | 55 |
| SWOG-9629 | Phase II | Chemo/Potentiator | First Line | 58 |
| SWOG-9629 | Phase II | Chemo/Potentiator | Second Line or More | 48 |
| SWOG-9924 | Phase II | Biologic | First Line | 53 |

**Table 2**

Trial characteristics for primary analysis cases. N=1,132.

| Protocol | N[&] | Activation Year | Type of Trial | Treatment | Progression Criteria | Assessment Interval |
|---|---|---|---|---|---|---|
| ECOG-1298 (18) | 30 | 1999 | Phase II | Gem/Docetaxel | ECOG[*] | 8 weeks |
| ECOG-2297 Arm 1 (19) | 156 | 1998 | Phase III | Gemcitabine | ECOG[*] | 8 weeks |
| ECOG-2297 Arm 2 (20) | 149 | 1998 | Phase III | Gem/5FU | ECOG[*] | 8 weeks |
| ECOG-3296 (20) | 36 | 1996 | Phase II | Gem/5FU | ECOG[*] | 8 weeks |
| NCCTG-0043 (21) | 40 | 2001 | Phase II | Gem/ISIS-2503 | RECIST[#] | 6 weeks |
| NCCTG-014C Arm B (22) | 36 | 2002 | Rand. Phase II | Gem/PS-341 | RECIST[#] | 6 weeks |
| NCCTG-034A (23) | 70 | 2005 | Phase II | Gem/Oxali/Bev | RECIST[#] | 2 cycles/2 mos. |
| NCCTG-984351 (24) | 54 | 1999 | Phase II | Gem/Oxaliplatin | WHO[$] | 6 weeks |
| SWOG-0205 Arm 1 (25) | 282 | 2004 | Phase III | Gem/Cetuximab | RECIST[#] | 2 cycles/8wks |
| SWOG-0205 Arm 2 (25) | 279 | 2004 | Phase III | Gemcitabine | RECIST[#] | 2 cycles/8wks |

[*] ECOG Solid Tumor Response Criteria (10)

[#] Response Evaluation Criteria In Solid Tumors (7)

[$] World Health Organization Guidelines (11)

[&] Number of cases included in the current analysis

**Table 3**

**Study population and published results for included trials**

| Protocol | N | M/F (%) | PS 0/1/2 (%) | PFS | OS |
|---|---|---|---|---|---|
| **ECOG-1298** | 32 | 50 / 50 | 12 / 72 / 16 | 2.1 Median PFS | 4.7 mos Median OS |
| **ECOG-2297 Arm 1** | 162 | 54 / 46 | 35 / 52 / 14 | 2.2 mos. Median PFS | 5.4 mos. Median OS |
| **ECOG-2297 Arm 2** | 160 | 52 / 48 | 23 / 64 / 14 | 3.4 mos Median PFS | 6.7 mos Median OS |
| **ECOG-3296** | 36 | 69 / 31 | 28 / 53 / 19 | 2.4 mos. Median TTF[*] | 4.3 mos Median OS |
| **NCCTG-0043** | 48 | 37 / 63 | 34 /58 / 8 | 3.8 mos. Median PFS | 6.7 mos. Median OS |
| **NCCTG-014C Arm B** | 36 | 53 / 47 | 35 / 58 / 7 | 2.4 mos. | 4. 8 mos. |
| **NCCTG-034A** | 82 | 70 / 27[&] | 40 / 54 /6[&] | 5.7 mos | 8.1 mos. |
| **NCCTG-984351** | 46 | 56 / 44 | 35/41 / 20 | 4.5 mos. | 6.2 mos. |
| **SWOG-0205 Arm 1** | 371 | 54 / 46 | 87 / 13[#] | 3.0 mos. | 5.9 mos. |
| **SWOG-0205 Arm 2** | 372 | 51 / 49 | 87 / 13[#] | 3.4 mos. | 6.3 mos. |

[*] Time to treatment failure

[#] PS 0-1 / 2

[&] Percentages generated from data, not published report

**Table 4**

Patient Characteristics and Statistics from univariate and multivariate Cox proportional hazards regression models for 1,132 patients. (N=1,116 cases complete for data for multivariate models.)

| | | | Overall Survival | | Progression Free Survival |
|---|---|---|---|---|---|
| Factor | N (%) | Comparison | Single Factor HR (P) N=1132 | Multivariate HR (P) N=1116 | Single Factor HR (P) N=1132 |
| **Sex** | | | | | |
| Female | 518 (46%) | | | | |
| Male | 614 (54%) | Male vs. Female | 1.16 (0.017) | 1.15 (0.02) | 1.09 (0.16) |
| **Performance Status** | | | | | |
| 0 | 316 (28%) | | | | |
| 1 | 650 (57%) | P.S. 1 vs. P.S. 0 | 1.28 (<0.0001) | 1.27 (<0.0006) | 1.12 (0.09) |
| 2 | 150 (13%) | P.S. 2 vs. P.S. 1 | 1.86 (<0.0001) | 1.87 (<0.0001) | 1.58 (<.0001) |
| Not reported | 16 (1%) | None | NA | NA | NA |
| **Race** | | | | | |
| Caucasian | 1020 (90%) | Cauc. vs. rest | 0.91 (0.33) | NA | 0.85 (0.10) |
| Black | 76 (7%) | Black vs. rest | 1.18 (0.16) | NA | 1.12 (0.35) |
| Asian/Other | 18 (2%) | Asian/oth. vs. rest | 0.95 (0.78) | NA | 1.29 (0.13) |
| Not reported | 18 (2%) | None | NA | NA | NA |
| **Bilirubin** | | | | | |
| Normal | 511 (48%) | | | NA | |
| >Normal | 81 (7%) | > Normal vs. Normal | 1.13 (0.046) | NA | 1.06 (0.38) |
| Not reported | 540 (47%) | None | | | |
| **SGOT/SGPT** | | | | | |
| Normal | 474 (42%) | | | | |
| >Normal | 315 (28%) | > Normal vs. Normal | 1.35 (<.001) | NA | 1.31 (<0.001) |
| Not reported | 343 (30%) | None | NA | NA | |

**Table 5**

Overall survival statistics for advanced pancreatic cancer according to sex and performance status.

| | | Female | | | | Male | | |
|---|---|---|---|---|---|---|---|---|
| | | 6 Month / 12 Month Overall Survival Rates | | **Median OS Months (95% CI)** | | 6 Month / 12 Month Overall Survival Rate | | **Median OS Months (95% CI)** |
| | **N** | **Predicted***  | **Observed** | | **N** | **Predicted*** | **Observed** | |
| P.S. 0 | 130 | 61% / 26% | 64% / 30% | 7.62 (6.41 – 9.4) | 158 | 60% / 20% | 57% / 17% | 6.65 (5.91 – 7.2) |
| P.S. 1 | 286 | 47% / 18% | 45% /16% | 5.37 (5.06 – 6.05) | 326 | 46% /14% | 47% / 16% | 5.73 (4.76 – 6.34) |
| P.S. 2 | 68 | 27% / 4.0% | 26% / 4.4% | 3.81 (2.66 – 4.86) | 78 | 26% / 2.9% | 27% 2.6% | 2.51 (1.71 - 3.25) |

*
Predicted rates are conditional on the levels of the covariates (sex and performance status) ad are derived from a logistic regression model.

**Table 6**

Progression free survival statistics for advanced pancreatic cancer according to sex and performance status.

| | | Female | | | | Male | | |
|---|---|---|---|---|---|---|---|---|
| | | **6 Month Progression-Free Survival Rate** | | **Median OS Months (95% CI)** | | **Month Progression-Free Survival Rate** | | **Median OS Months (95% CI)** |
| | **N** | **Predicted**[*] | **Observed** | | **N** | **Predicted**[*] | **Observed** | |
| P.S. 0 | 130 | 26% | 27%± | 3.60 (2.66 – 4.73) | 158 | 25% | 23%± | 3.63 (2.83-3.94) |
| P.S. 1 | 286 | 25% | 24%± | 3.45 (2.66-3.68) | 326 | 24% | 24%± | 2.51 (2.07-3.38) |
| P.S. 2 | 68 | 15% | 13%± | 2.14 (1.58-3.15) | 78 | 14% | 15%± | 1.64 (1.41-2.76) |

[*] Predicted rates are conditional on the levels of the covariates (sex and performance status) ad are derived from a logistic regression model.