DOI:10.1093/jnci/dju002
First published online March 13, 2014

ARTICLE

Comparison of Survival Outcomes Among Cancer Patients Treated In and Out of Clinical Trials

Joseph M. Unger, William E. Barlow, Diane P. Martin, Scott D. Ramsey, Michael LeBlanc, Ruth Etzioni, Dawn L. Hershman

Manuscript received February 7, 2013; revised October 29, 2013; accepted December 18, 2013.

Correspondence to: Joseph M. Unger, PhD, MS, SWOG Statistical Center, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, M3-C102, PO Box 19024, Seattle, WA 98109-1024 (e-mail: junger@fhcrc.org).

Background Clinical trials test the efficacy of a treatment in a select patient population. We examined whether cancer clinical trial patients were similar to nontrial, "real-world" patients with respect to presenting characteristics and survival.

Methods We reviewed the SWOG national clinical trials consortium database to identify candidate trials. Demographic factors, stage, and overall survival for patients in the standard arms were compared with nontrial control subjects selected from the Surveillance, Epidemiology, and End Results program. Multivariable survival analyses using Cox regression were conducted. The survival functions from aggregate data across all studies were compared separately by prognosis (≥50% vs <50% average 2-year survival). All statistical tests were two-sided.

Results We analyzed 21 SWOG studies (11 good prognosis and 10 poor prognosis) comprising 5190 patients enrolled from 1987 to 2007. Trial patients were younger than nontrial patients ($P < .001$). In multivariable analysis, trial participation was not associated with improved overall survival for all 11 good-prognosis studies but was associated with better survival for nine of 10 poor-prognosis studies ($P < .001$). The impact of trial participation on overall survival endured for only 1 year.

Conclusions Trial participation was associated with better survival in the first year after diagnosis, likely because of eligibility criteria that excluded higher comorbidity patients from trials. Similar survival patterns between trial and nontrial patients after the first year suggest that trial standard arm outcomes are generalizable over the long term and may improve confidence that trial treatment effects will translate to the real-world setting. Reducing eligibility criteria would improve access to clinical trials.

JNCI J Natl Cancer Inst (2014) 106(3): dju002 doi:10.1093/jnci/dju002

Randomized cancer clinical trials represent a final step in evaluating the efficacy of new treatments. However, few adult cancer patients participate in trials (<3%) in the United States (1,2). Reasons for low rates of clinical trial participation are numerous (3–5). Trials may not be available for patients willing to participate, or when they are available, patients are often excluded because they do not meet trial eligibility criteria (6–9).

Trial eligibility criteria must satisfy two opposing factors (10). They must be sufficiently narrow to establish a homogeneous sample, so the effect of treatment is roughly consistent across the cohort. Eligibility criteria that are too broad risk including patients for which the treatment is not optimal, which could mask the overall treatment effect. Eligibility should also be sufficiently broad that the results are generalizable. One possible difference between trial and nontrial patients is that trial eligibility criteria rule out poor-prognosis patients with prior comorbid conditions. Yet if the trial cohort is otherwise representative of the general cancer population with respect to cancer histology and stage, any differences in survival induced by ruling out poor-prognosis patients may not endure over time.

Despite attempts by clinical trialists to establish equipoise between homogeneity and generalizability, clinical trials are sometimes criticized for sacrificing generalizability (11). To assess generalizability in a systematic fashion, we evaluated whether presenting characteristics and survival outcomes for patients on the standard arms of a series of randomized phase III cancer clinical trials were representative of outcomes in patients receiving non–clinical trial treatment.

Methods

Cancer clinical trial data were from SWOG, a national clinical trials consortium sponsored by the National Cancer Institute. Nontrial data were from the Surveillance, Epidemiology, and End Results (SEER) cancer registry (12).

We conducted an analysis of randomized phase III studies from the SWOG historical database over a 25-year period (1987–2011). SWOG studies must have been published and must have had upfront randomization because studies with postregistration filtering of patients before receipt of standard treatment could not be

reproduced using SEER. SWOG studies of recurrent disease were excluded (because SEER indexes reported case patients according to first diagnosis of a unique tumor type), as were studies with nonsurvival endpoints. Figure 1 shows how approximately two-thirds of candidate trials were excluded for these reasons.

To be included, data to replicate the essential primary site, histology, and stage specifications from the SWOG study must have been available in SEER. Staging criteria included both TNM staging and, where appropriate, surgical and nodal staging. Studies that relied on tumor characteristics not available in SEER were excluded. We excluded positive SWOG studies for which there was also a trend toward improved survival over time in the corresponding SEER population because the standard arms for these SWOG studies likely no longer reflected community standard care at study completion. Only subjects on the standard arm were included, and corresponding SEER patients must have had a diagnosis date during the SWOG study's enrollment period. Assuming the SWOG standard arm represented standard-of-care in the general cancer population during the study enrollment period, this allowed comparison between trial and nontrial patients with approximately similar treatments. The age limits specified in SWOG study eligibility were applied to the corresponding SEER datasets. Nearly all SWOG studies excluded patients with prior malignancies; for comparability, only SEER patients with first primaries were included.

Statistical Considerations

Comparisons between SWOG and SEER patients with respect to age (<65 years vs ≥65 years), sex, race (black vs white vs other), and stage were conducted across the panel of SWOG studies. For studies with more than one stage, stage was dichotomized into approximately equal groups to enable a consistent method of adjustment across the different studies. To test whether there was a global trend in stage or demographic rates across the panel of studies, the study-specific rates for both SEER and SWOG were converted to z scores (one for each study), and a one-sample t test was conducted on the difference in the z scores between SEER and SWOG.

For each study, Kaplan–Meier plots were generated to explore patterns of survival between SEER and SWOG patients, and Cox regression was used to estimate the hazard ratio and 95% confidence interval (CI) for the impact of trial participation, accounting for age, sex, race, stage, and year of enrollment (13,14). Studies were categorized as good (≥50%) vs poor (<50%) prognosis based on observed results using average 2-year Kaplan–Meier survival estimates.

To further explore differences in survival patterns, SWOG and, separately, SEER patients were combined by prognosis. To construct an equally weighted sample, 50 patients from each SWOG study and each corresponding SEER cohort were randomly selected. This process was averaged across 1000 repeat random samples. Kaplan–Meier plots and corresponding smoothed hazard functions (using Kernel-based methods) of the aggregate datasets were examined (15–17).

Based on the patterns observed using smoothed hazard function analysis, we applied landmark survival analysis to assess survival patterns related to trial participation given survival of the patient for a certain duration.

The contributions of cause of death to survival patterns were also investigated. SEER codes cause of death according to the *International Classification of Diseases, Tenth Edition*. In SWOG, a death was deemed cancer related if it followed a documented cancer progression. SWOG rates were adjusted using cause-of-death data available for a subset of patients (see Supplementary Methods, available online).

Finally, we assessed the extent to which study factors determined variation in survival outcomes. We estimated components of variation of the factors by comparing the partial log-likelihoods from nested models. We took the average of both forward and backward nesting approaches, with factors rank-ordered for model inclusion according to their χ^2 statistic in a multivariable model.

All analyses were limited to survival in the first 5 years after diagnosis to emphasize outcomes related to cancer and its treatments. All statistical tests were two-sided.

Results

Study Selection

Of 102 SWOG studies examined, 64 were initially excluded (Figure 1). Seventeen of the remaining 38 studies were excluded because of inadequate SEER data on essential tumor characteristics.

Study Profiles and Eligibility

Twenty-one studies (n = 21/38; 55%) met the specified study inclusion criteria (Table 1) (18–38). The study sample included both early- and late-stage cancers from many cancer types. A total of 5190 SWOG patients and 69 187 SEER patients from 1987 to 2007 were analyzed.

Table 2 summarizes additional eligibility criteria from the SWOG study that do not specifically pertain to histology or tumor characteristics. Nearly all studies had prior systemic therapy exclusions and required adequate kidney, liver, and hematologic function. The majority of studies required no current evidence or history of cardiac dysfunction. Other common exclusion criteria included other serious medical conditions, diseases, or active infections, and low patient functional status. Most of the criteria in Table 2 could not be accounted for using SEER data.

The mean total number of eligibility criteria for a given study was 16.1, of which 9.8 (60%) were related to comorbidity or performance status.

Demographic Factors and Stage

Figure 2 shows the difference between SEER and SWOG patients for each demographic and stage factor. The SEER cohort was consistently more likely to be older and, to a lesser degree, female, but there were no panel-wide trends in the proportion of patients with higher stage or black race.

Overall Survival Comparisons Between SWOG and SEER

Figure 3 shows both unadjusted and multivariable (adjusted) hazard ratios comparing overall survival between SWOG and SEER cohorts in descending order of average 2-year survival. Eleven studies had average 2-year survival of 50% or greater (good prognosis) and 10 studies had 2-year survival of less than 50% (poor prognosis). For none of the good-prognosis studies did survival for SWOG patients statistically significantly differ from survival for SEER patients in multivariable analysis, whereas for nine of

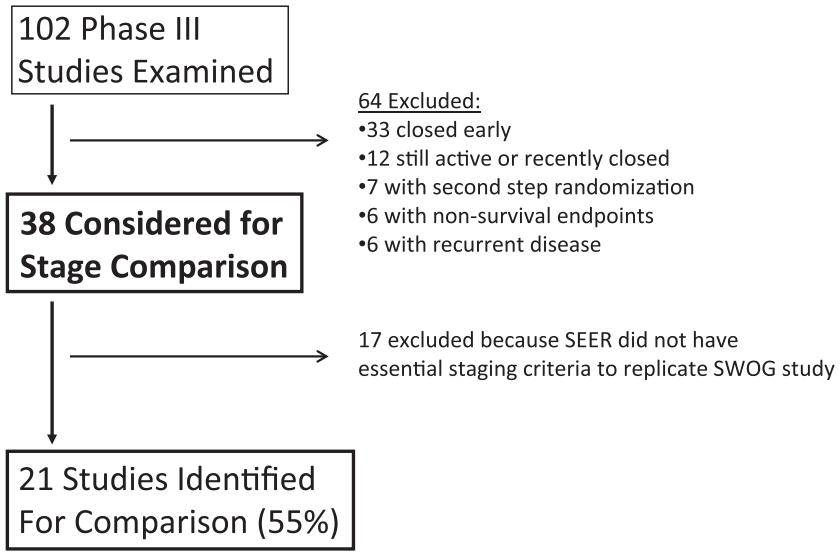

Figure 1. Study identification flow diagram. One hundred two phase III SWOG studies were examined over the 25-year period from 1987 to 2011. Among these, 64 were excluded from further consideration, 33 because of early closure (of which 30 were closed early because of poor accrual, two were closed early because of changed relationship with the drug manufacturer, and one was a positive study based on progression-free survival), 12 were still active or recently closed, seven did not have upfront randomization, six had nonsurvival endpoints, and six were studies for recurrent disease. Of the 38 considered for comparison with Surveillance, Epidemiology, and End Results (SEER) registry data, 17 were excluded because SEER did not have essential staging criteria to replicate the SWOG study. In the end, 21 of 38 studies (55% of those considered for stage comparison) were identified.

10 poor prognosis studies, SWOG patients had statistically significantly lower risk of death ($P < .001$).

We found no evidence that the hazard ratios for trial participation differed over calendar time for either good-prognosis ($P = .50$) or poor-prognosis ($P = .69$) studies (see Supplementary Figure 1, available online). Also, results did not substantively change when a covariable for Hispanic ethnicity was added to the multivariable models (see Supplementary Methods, available online).

Differences in Aggregate Survival Patterns Between SWOG and SEER Patients

Examination of the individual study-specific survival curves (see Supplementary Figure 2, available online) indicated a frequent pattern of an early survival advantage for SWOG patients that waned over time for both good- and poor-prognosis cancers. Using aggregate data, we examined Kaplan–Meier plots of overall survival and corresponding smoothed hazard functions (Figure 4). For both good- and poor-prognosis patients, the hazard function for SWOG patients was initially much lower than the hazard function for SEER patients. But, by year 1, the hazard functions for both SEER and SWOG patients no longer differed, suggesting that trial participation was associated with better survival only in the first year. Importantly, this analysis revealed a consistent association of trial participation and survival that was not evident in the individual survival analyses for good-prognosis studies, likely because of limited power in that setting.

Average Effect Accounting for the First-Year Survival Difference

For good-prognosis patients, the mean of the adjusted hazard ratios for overall survival comparing SWOG with SEER patients shown in Figure 3 was not statistically different from 1.0 (mean = 0.96; 95% CI = 0.92 to 1.01; $P = .12$). We analyzed the subset of patients who survived 1 year using landmark survival analysis. The results were similar (mean = 1.05; 95% CI = 0.96 to 1.14; $P = .22$). However, for poor-prognosis patients, the mean of the multivariable hazard ratios shown in Figure 3 was much less than 1.0 (mean = 0.74; 95% CI = 0.64 to 0.84; $P < .001$). Conditioning on 1-year survival, this difference was no longer evident (mean = 1.05; 95% CI = 0.95 to 1.15; $P = .27$), reinforcing the observation that the impact of trial participation endured for only about 1 year.

Analysis of Cancer-Specific and Non-Cancer-Specific Events

The proportion of patients experiencing cancer-related and non-cancer-related deaths relative to the number of patients at risk was analyzed by year. Non-cancer-related deaths were lower in SWOG patients, although this difference was small and relatively stable across all 5 years of follow-up (Figure 5). In contrast, cancer-related deaths were notably lower in the first year in SWOG patients but similar to SEER patients in later years. Therefore the difference in the patterns of death for trial vs nontrial patients between year 1 vs years 2 to 5 is largely attributable to different patterns of cancer-related deaths.

Attributable Variation

In the non-sex-specific studies, disease and stage explained 92.2% of the relative variation in survival outcomes, followed by age (5.2%), trial participation (1.5%), race (0.6%), and sex (0.5%). In the first year only, estimate of variation in survival outcomes attributable to disease and stage was 88.4% and to trial participation was 4.9%, compared with 92.7% and 1.2%, respectively, after 1 year.

Table 1. SWOG studies included for comparison with the Surveillance, Epidemiology, and End Results (SEER) registry*

	SWOG criteria				Corresponding SEER criteria		
Cancer and study no.	**Years of accrual**	**Histology**	**Major tumor characteristic criteria from SWOG studies†**	**SWOG No.**	**SEER No.**	**ICD-O-3 primary site**	**Histology code**
Brain S0001	2001–2005	Glioblastoma multiforme/ gliosarcoma	Biopsy or surgical resection prior to registration	89	2264	C710–725	9440–9444
Breast S9313	1994–1997	Adenocarcinoma‡	Stage T1-3, N0, M0 (selected stages I-III; no locally advanced disease) Axillary dissection required ≥6 nodes removed and examined ≤3 positive nodes Tumor >2 cm and ER/PR (-) or (+); or, 1-3 (+) axillary nodes Prior mastectomy or breast sparing surgery	1423	9941	C500–509	8500–8530
Breast S0012	2001–2005	Locally advanced or inflammatory breast carcinoma	Stage IIB–IIIB (M0)	391	2855	C500–506, C508–509	Any
GI-Gastric S9008	1991–1998	Adenocarcinoma§	Stage IB–IV (M0) Prior en bloc surgery	283	2487	C150–155, 58–66, 68–69	8140–8800
GI-Pancreas S0205	2004–2006	Adenocarcinoma‖	Locally advanced (not surgically resectable, ie, no prior surgery) or metastatic disease	82	1943	C250–254, C257–259	8140
GU-Bladder S8710	1988–1997	Transitional cell carcinoma	Stage T2–T4A (no metastasis)	148	2377	C670–679	8120–8124
GU-Bladder S8795	1988–1992	Transitional cell carcinoma (including papillary)	Stage Ta–T1 and grade I–IV Completely resected	191	5059	C670–679	8120–8124, 8130
GU-Prostate S8894	1989–1994	Adenocarcinoma	Stage D2	534	5961	C619	8140
GU-Renal S8949	1991–1998	Carcinoma	Metastatic No nephrectomy (standard arm)	95	1569	C649	8312
GYN-Cervix S8797	1990–1996	Squamous cell carcinoma, adenocarcinoma, or adenosquamous carcinoma	Stages IA2, IB, or IIA Radical hysterectomy with total pelvic lymphadenectomy Positive pelvic or parametrial, and negative para-aortic, nodal involvement	130	137	C530–531, C538–539	8070–8, 8140–7, 8260–3, 8310–84, 8560–62
LEUK-AML S9031	1991–1994	AML	FAB classes M0–M2, M4–M7 (excluded M3s beginning in August, 1992)	85	1672	C420–1, C424	9801, 9840, 9861, 9866–7, 9871–74, 9891, 9896, 9910
LEUK-AML S9333	1995–1998	AML	FAB classes M0–M2, M4–M7 (excluded M3s)	129	2320	C420–1, C424	9801, 9840, 9861, 9867, 9871–74, 9891, 9896, 9910
Lung-NSCLC S8738	1988–1990	Squamous cell carcinoma, adenocarcinoma, and large cell carcinoma	M1 disease (including lung metastasis). Exclude patients with mets only to ipsilateral hilar nodes (N1) and/or mediastinal nodes (N2) or supraclavicular nodes (N3) ONLY	94	4084	C340–3, C348–9	8012, 8070–78, 8140–47
Lung-NSCLC S9308	1993–1995	Any NSCLC	Stage IIIB (based on positive pleural effusions or ipsilateral lung involvement) or stage IV	178	4755	C340–3, C348–9	8012, 8046, 8070–8, 8140–7, 8240–50, 8560, 9050–3

(Table continues)

Table 1 (Continued).

SWOG criteria					Corresponding SEER criteria		
Cancer and study no.	**Years of accrual**	**Histology**	**Major tumor characteristic criteria from SWOG studies†**	**SWOG No.**	**SEER No.**	**ICD-O-3 primary site**	**Histology code**
Lung-NSCLC S9509	1996–1997	Any NSCLC (except bronchioalveolar)	Stage IIIB with either 1) T4 disease due to malignant pleural effusion; 2) multiple lesions in a single lobe containing a T3 or T4 primary; or 3) lesions in multiple lobes of the ipsilateral lung for which one such lesion is T3 or T4;¶ or stage IV	205	4817	C340–3, C348–9	8012, 8046, 8070–8, 8140–7, 8240–9, 8560, 9050–3
Lung-NSCLC S9900	1999–2004	Any NSCLC	Selected stages IB (T2N0), II (T1–2, N1; or T3N0), or IIIA (T3N1) Limited to surgery type specified in protocol: lobectomy, sleeve resection, bilobectomy, or pneumonectomy (excludes limited resection or NOS)	168	829	C340–3, C348–9	8012, 8046, 8070–8, 8140–7, 8240–50, 8560, 9050–3
Lung-NSCLC S0003#	2000–2002	Squamous, adeno-, large cell, or NSCLC carcinoma	Use newly diagnosed, selected stage IIIB (based on positive pleural effusions) or stage IV	165	7727	C340–3, C348–9	8012, 8046, 8070–8, 8140–7
Lung-SCLC S0124	2002–2007	Any SCLC	Extensive disease	266	2790	C340–3, C348–9	8041–5
Melanoma S8642	1987–1990	Any melanoma	Stage II (thickness ≥1.5, N0, M0) or III (any T, N1-2, M0) Complete wide-excision of tumor (≥1 cm margin)**	96	738	C440–9	8720–72
Melanoma S9035	1992–1996	Any melanoma	Stage T3N0M0 (thickness 1.51–4.00 mm or Clark IV if thickness unknown) Complete wide-excision of tumor (≥1 cm margin)**	299	1347	C440–9	8720–72
Myeloma S8624	1987–1990	Multiple myeloma	Previously untreated	139	3515	C421	9732
TOTAL 21 studies	21 years (1987–2007)			5190	69 187		

* AML = acute myeloid leukemia; ER = estrogen receptor; FAB = French-American-British; GI = gastrointestinal; GU = genitourinary; GYN = gynecologic; LEUK = leukemia; NSCLC = non–small cell lung cancer; PR = progesterone receptor; SCLC = small cell lung cancer.

† All criteria listed in the table were explicitly accounted for in SEER. Additional tumor characteristic criteria that could not be accounted for explicitly in SEER include: Brain, S0001) Patients with three or more noncontiguous sites are ineligible; GI-Gastric, S9008) No ascites; no peritoneal seeding; no liver metastases or extra-abdominal metastases; GU-Bladder, S8710) One or more kidney and proximal ureter free of tumor and all other disease resectable; GU-Bladder, S8795) No recurrent tumor on cystoscopy within 4 weeks if first TURBT more than 4 weeks before registration; and, random biopsy or a negative urinary cytology; GU-Renal, S8949) Primary cancer must be amenable to surgery if patient did not otherwise have metastatic disease; Leukemia-AML, S9031 and S9333) Exclude blastic transformation of chronic myelogenous leukemia; Lung-NSCLC, S9509) Exclude stage IIIB tumors involving the superior sulcus; Lung-NSCLC, S9900) No patients with symptomatic tumors (T3,N0–N1) involving the superior sulcus; Melanoma, S9035) Lymphadenectomy must have resolved; patients with suspicious nodes must have regional lymph node dissection with negative nodes; Myeloma, S8642) Specific protein criteria; and, patients with immunoglobulin M myeloma not eligible.

‡ Excluding tubular, mucinous, papillary, sarcoma, lymphoma, apocrine, adenocystic, or squamous cell carcinoma; ductal or lobular carcinoma in situ allowed if one to three positive nodes. Patients with tumor greater than 1 cm and ER/PR(-) excluded from both SWOG and SEER datasets because of lack of ER/PR data in SEER during the study period.

§ Stomach and esophagogastric junction.

‖ Exclude endocrine tumors, lymphoma of pancreas, or ampullary cancer.

¶ For IIIB definition in SEER, simplified as IIIB with T3 or T4 extent-of-disease.

Although S0003 allowed recurrent patients, these were excluded. Comparison with SEER relied on newly diagnosed patients only.

** Detailed surgical resection criteria were specified.

Table 2. Eligibility criteria for SWOG clinical trials*

Cancer type and study No.	Min (max) age	No prior cancer†	Prior treatment exclusions	Organ function criteria‡	Max PS§	Other PS	Pregnant/ contra-ception	Serious medical condition‖	HIV	Prior TX timing	Other on-study therapy	No brain mets	Study drug allergy	Scan timing	No. other criteria¶
Brain, S0001	18	X	Ch, RT	K, P	2	—	X	X	X	—	X	—	X	X	0
Breast, S9313	NS	—	Ch, R, S	K, L, H, C	—	LTFU	X	X	—	X	—	—	—	X	1
Breast, S0012	NS	X	Ch, Hr, RT, S	K, L, H, C	2	—	X	—	X	—	X	—	—	X	0
GI-Gastric, S9008	NS	X	Ch, B, RT	K, L, H	2	—	X	X	—	X	—	—	—	X	1
GI-Pancreas, S0205	NS	X	Ch	K, L, H, C	2	—	X	—	X	X	X	X	—	X	0
GU-Bladder, S8710	NS	X	RT#	K, L, H, C	1	CURE	X	X	—	X	—	—	—	X	1
GU-Bladder, S8795	NS	X	Ch	K, L, H	2	LE	X	—	—	X	X	—	—	X	1
GU-Prostate, S8894	NS	X	Ch, Hr, B	K, L, H	3	—	—	X	—	—	X	—	—	X	0
GU-Renal, S8949	NS	X	Ch, Hr, B, RT**	K, L, H, C	2	—	X	—	—	—	X	X	—	X	0
GYN-Cervix, S8797	NS	X	Ch, Hr, B, RT#	K, L, H	2	—	—	X	—	X	—	—	—	X	1
LEUK-AML, S9031	56	X	Ch	K, L, C	3	—	X	—	—	—	—	—	—	X	0
LEUK-AML, S9333	56	X	Ch	K, L, C	3	—	X	—	—	—	—	—	—	X	1
Lung-NSCLC, S8738	NS	X	Ch	K, H, C	2	LE	X	—	—	—	X	X	—	X	0
Lung-NSCLC, S9308	18	X	Ch, B	K, L, H	1	—	X	X	—	X	—	X	—	X	1
Lung-NSCLC, S9509	18	X	Ch, B	K, L, H, C	1	—	X	X	—	—	—	X	X	X	1
Lung-NSCLC, S9900	18	X	Ch, RT	K, L, H, P	1	—	X	X	—	—	X	—	X	X	1
Lung-NSCLC, S0003	NS	X	Ch, B	K, L, H	1	—	X	—	—	X	—	X	X	X	0
Lung-SCLC, S0124	18	X	Ch, RT††	K, L, H	1	—	X	—	X	X	—	—	—	X	1
Melanoma, S8642	18 (70)	X	Ch, Hr, B, RT	K, L, H, C	1	—	X	X	—	—	X	—	—	X	2
Melanoma, S9035	18	X	Ch, Hr, B, RT	K, L, H, C	1	—	X	—	—	X	X	—	—	X	0
Myeloma, S8624	NS	X	Ch	H‡‡, C	3	—	—	X	—	—	—	—	—	X	1

* Only the first two criteria listed (age and prior cancer) were explicitly accounted for in the Surveillance, Epidemiology, and End Results (SEER) registry. All other criteria could not be accounted for based on SEER data. Eligibility criteria that related to comorbidity or performance status included prior treatment exclusions, prior malignancy exclusions, performance status, organ function status, human immunodeficiency virus status, serious medical conditions, brain metastases, study drug allergy, and maximum age limit. Empty cells (cells with dashes) indicate the particular eligibility criterion was not included in the study protocol.

AML = acute myeloid leukemia; B = biologic therapy; C = cardiac; Ch = chemotherapy; CURE = potentially curable; GI = gastrointestinal; GU = genitourinary; GYN = gynecologic; H = hematologic; HIV = human immunodeficiency virus; Hr = hormonal therapy; K = kidney; L = liver; LE = minimum life expectancy; LEUK = leukemia; LTFU = adequate health for long-term follow-up; NS = not specified; NSCLC = non–small cell lung cancer; P = pulmonary; PS = performance status; RT = radiation therapy; S = surgery; SCLC = small cell lung cancer; TX = treatment.

† Typically requires no prior malignancy except adequately treated non-melanoma skin cancer, in situ cervical cancer, or other cancer for which the patient has been disease free for 5 or more years.

‡ Organ function criteria were based primarily on the following tests: for kidney, creatinine clearance and/or serum creatinine; for liver bilirubin, serum glutamic oxaloacetic transaminase and/or serum glutamic pyruvate transaminase; and for hematologic, white blood count and platelets.

§ Performance status is a measure of the patient's well-being and activity level. In SWOG, the coding scheme is: 0 = asymptomatic or fully active; 1 = symptomatic but completely ambulatory; 2 = symptomatic but in bed less than 50% of day; 3 = symptomatic, more than 50% of time in bed, but not bedbound; 4 = completely disabled or bedbound.

‖ Including active infections.

¶ Other eligibility criteria include: Breast, S9313) Patients with breast-sparing surgery must plan RT after chemotherapy; GI-Gastric, S9008) Good caloric intake of 1500 or more calories/day required; GU-Bladder, S8710) Normal organ function required; GU-Bladder, S8795) Must be at increased risk of papillary tumor recurrence; GYN-Cervix, S8797) No pelvic inflammatory disease; Leukemia-AML, S9333) Exclude if marrow unaspirable and white blood cells and blasts + promyelocytes + promonocytes outside normal limits; Lung-NSCLC, S9308) No grade 2 or greater neuropathy; Lung-NSCLC, S9509) No grade 2 or greater neuropathy; S9900) No grade 2 or greater neuropathy; Lung-SCLC, S0124) Prior brain metastases must have been treated; Melanoma, S8642) No known seizure disorder or known central nervous system disease; no prior organ transplant; Myeloma, S8624) Patients must have objective evidence, or be symptomatic from, AML.

Pelvic.

** Except palliative.

†† Except brain.

‡‡ Based on M-component.

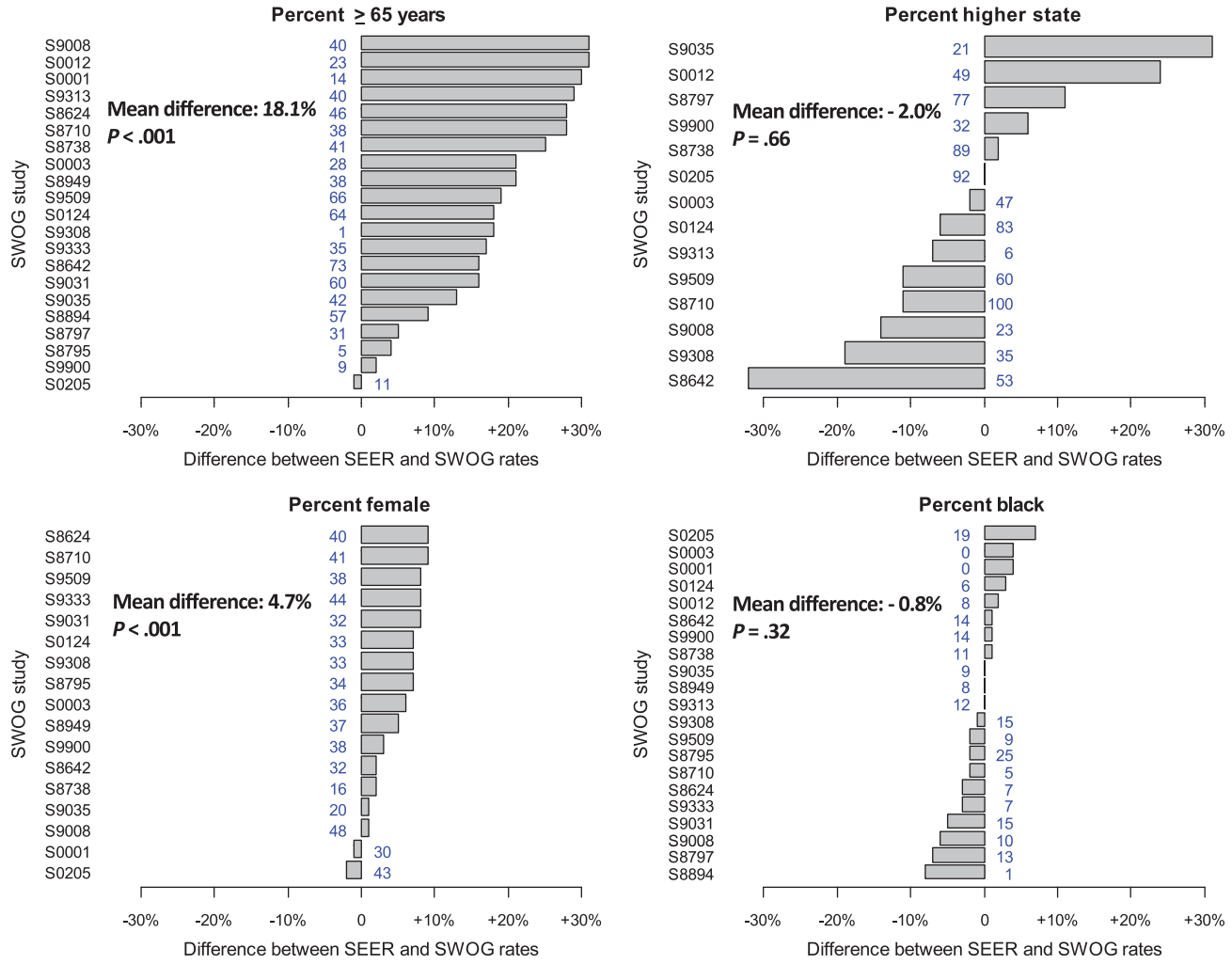

Figure 2. Horizontal barplots of the difference between Surveillance, Epidemiology, and End Results (SEER) and SWOG patients for each demographic and stage factor, in descending order of the absolute difference in percentages between SWOG and SEER cohorts. The SWOG percentage is also shown in each figure. **Bars to the right of center** indicate a higher proportion in SEER, and **bars to the left of center** indicate a higher proportion in SWOG.

Discussion

We found that trial participation was associated with better survival only in the first year. Short-term estimates of absolute survival probabilities from clinical trials may be optimistic (Figure 4). Physicians who use clinical trial results to assist in making treatment decisions should be aware of this phenomenon. Better short-term survival for trial patients is likely related to the exclusion of sicker patients from trials through eligibility criteria pertaining to comorbidity and performance status. These exclusions also resulted in trial cohorts that were much younger and somewhat less likely to be female, consistent with prior reports (39,40).

We did not explicitly assess whether the treatment effect in a clinical trial translates (ie, generalizes) to the broader cancer population. Such a study would require a comparison between experimental and standard arm treatments occurring in the general cancer population at the same time as the clinical trial is conducted. However, similar standard arm outcomes beyond the first year may improve confidence that efficacy of treatment in a trial translates to the real-world setting. This conclusion relies on the assumption that trial participation would impact standard and experimental treatment arms similarly and would not apply in instances where new treatments have too much toxicity or poor compliance. Importantly, we found no evidence that the association of trial participation and survival increased over calendar time, which might be expected if new treatments adopted into standard care do not show the same benefit as observed in the clinical trial. This suggests that most patients may also benefit from the new treatments, even if not participating in trials.

The most reliable way to establish the causal relationship between trial participation and outcome would be to randomize patients to be offered a clinical trial vs not offered a clinical trial (41) Such a study would be practically and ethically difficult. Instead, the literature is based on observational studies, which focus on presenting characteristics and absolute survival differences between trial and nontrial patients. Identification of the appropriate nontrial control group is crucial to inference because any observational design will be limited by unmeasured confounding, whether trial patients are compared with eligible nontrial control subjects (bias with respect to factors associated with refusing trial participation),

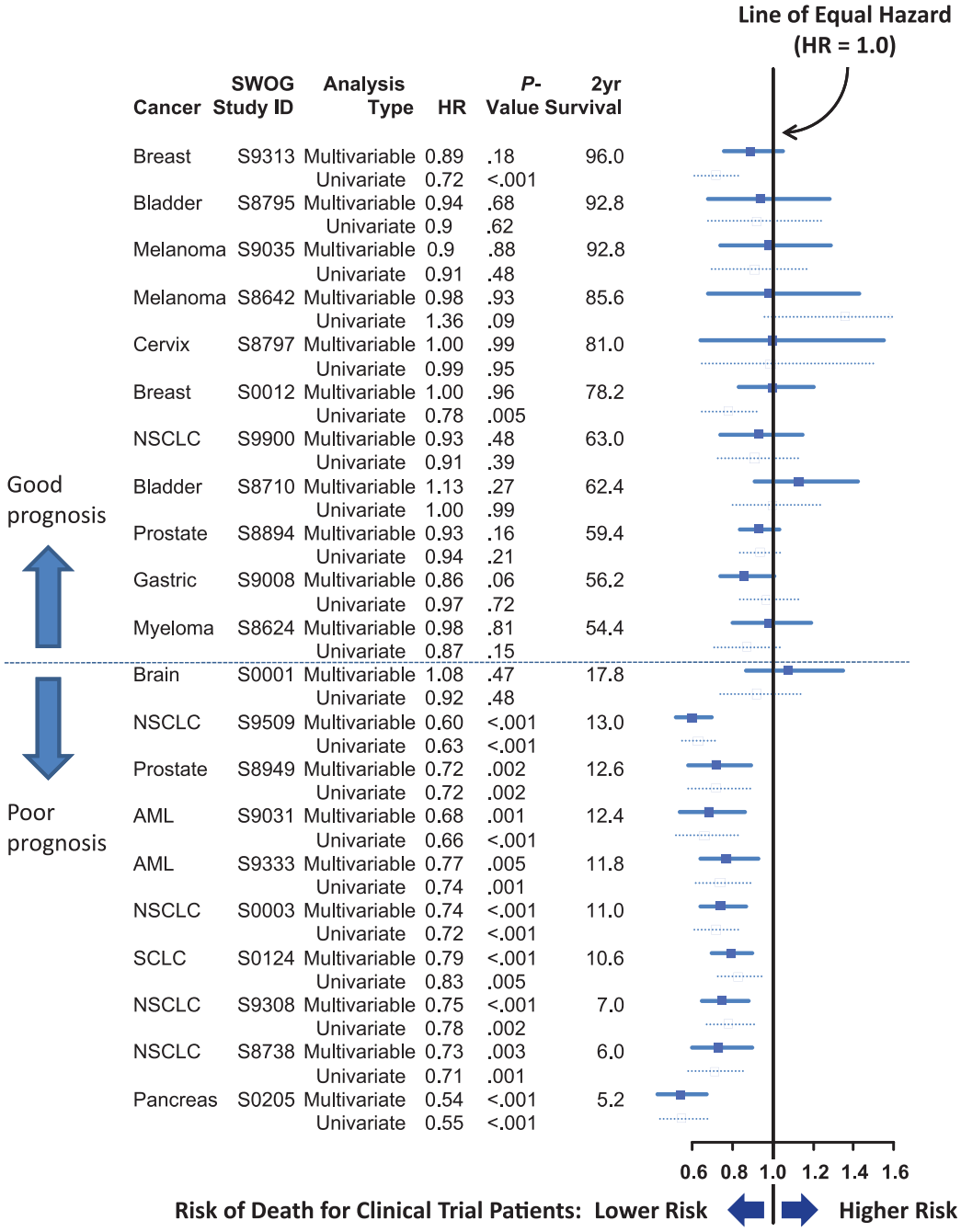

Figure 3. Forest plot of univariate and multivariable hazard ratios (HRs) for overall survival, by study, ordered in descending order of average 2-year overall survival. In univariate analyses, two of 11 (18%) good-prognosis studies and nine of 10 (90%) poor-prognosis studies showed evidence of a survival benefit for trial patients ($P = .002$ by Fisher exact test). In multivariable analyses, zero of 11 good-prognosis studies and nine of 10 poor-prognosis studies showed evidence of a survival benefit for trial patients ($P < .001$). AML = acute myeloid leukemia; NSCLC = non–small cell lung cancer; SCLC = small cell lung cancer.

ineligible control subjects (bias with respect to prognosis), or population control subjects (multiple biases) (41). These studies most often focused on single trial vs nontrial comparisons, raising the issue of subjective study selection.

Both Peppercorn et al. (41) and Edwards et al. (42) reviewed the historical literature. Both found that a majority of comparisons from cancer studies showed evidence of better outcomes for trial patients, with no evidence of harm. Peppercorn et al. (41) concluded that there was no strong evidence of a benefit for trial patients, in part because of methodological issues with the nontrial comparator groups, whereas Edwards et al. (42) concluded that there was positive, albeit weak, evidence that participation in trials improves outcomes. Other reviews and studies also found mixed evidence (43–46).

The inconclusive picture offered by the literature could be related to the transient impact of trial participation on survival found in this study. We re-examined the cancer studies included in two prior reviews (41,42). Studies were categorized as good or poor prognosis as defined in this study. In total, there were 36 comparisons from 27 studies (see Table 3)

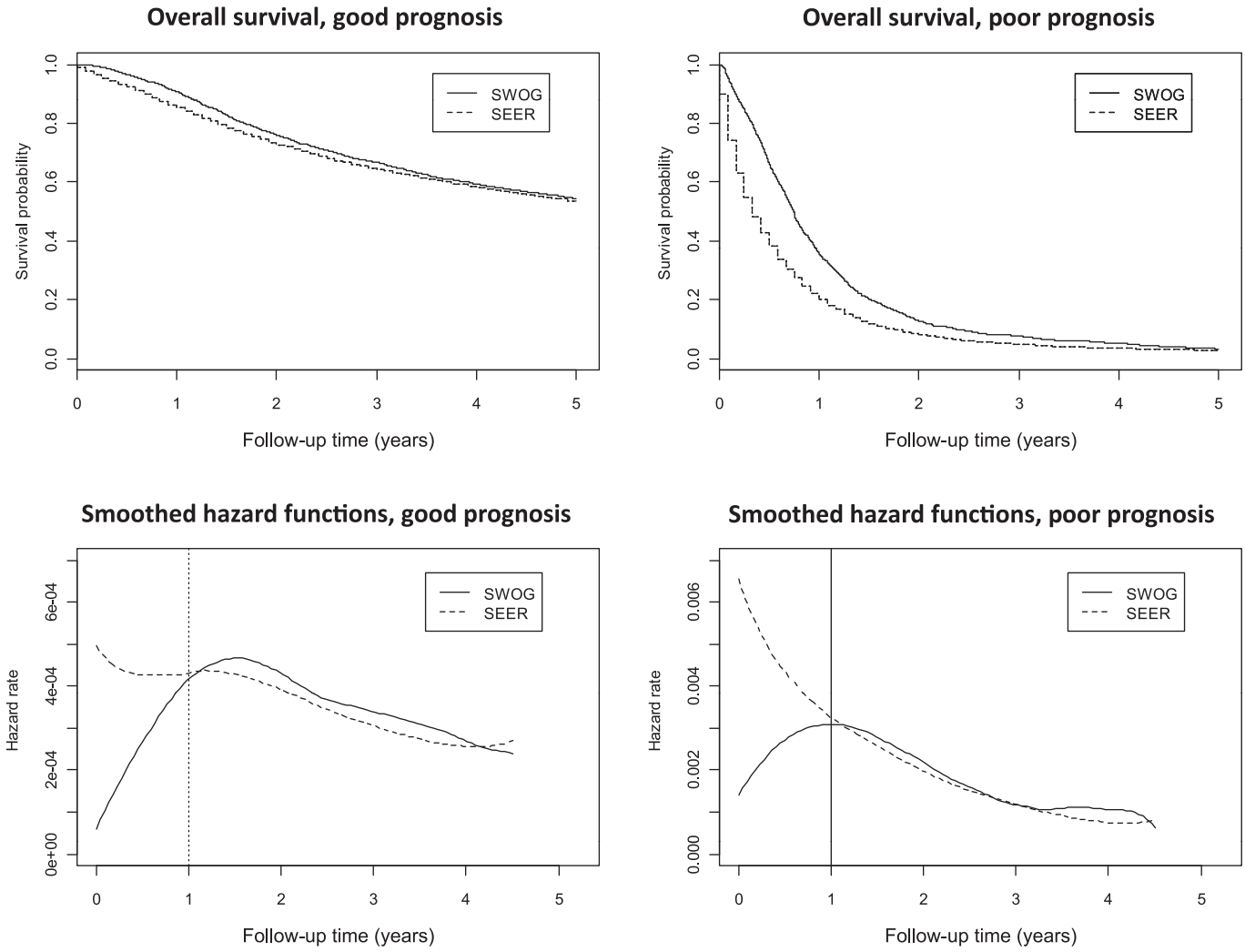

Figure 4. Overall survival and corresponding hazard functions for aggregate (equally weighted) study data by prognosis.

(47–73). Fifty-six percent of good-prognosis studies showed evidence of survival benefit for trial patients, compared with 82% of poor-prognosis studies, a pattern consistent with but not as extreme as the pattern found in this study. A similar pattern was found among comparisons that included multivariable analyses only (47,49,51–58,60–62,67–70) and adult cancers only (47–57,60,61,64–66,72,73).

We compared trial vs nontrial patients who were similar with respect to histology, stage, age, de novo presentation, year of diagnosis, race, sex, and treatment. What remained were differences between databases that we could not account for. Trial patients could benefit from changes in behavior or outlook associated with being under observation (the "Hawthorne" effect) (74) or from care that is administered according to strict protocol (75). Alternatively, none of the eligibility criteria outlined in Table 2, the majority of which pertain to performance status and comorbidities, could be accounted for. Therefore trial patients likely exhibit better outcomes because eligibility criteria prevent sicker patients from enrolling on study. These enrollment restrictions appear to primarily limit early cancer deaths, suggesting that comorbidity and performance status identify residual variation in cancer-specific survival even after accounting for stage. Unfortunately, the extent to which the survival differences were related to patient selection or other factors cannot be estimated with these data.

This study also had some limitations. We were unable to account for the actual treatments of the nontrial control patients. It is inevitable that not all nontrial patients in SEER received standard of care for their histology and stage and may have received no treatment. The use of different databases with different methods of data collection may induce different patterns of endpoint assessment, which could impact analyses of cancer-specific events in particular. Further, SEER patients have been shown to have, on average, higher socioeconomic status; thus SEER data are not precisely representative of the US cancer population (76,77). Because trial patients also tend to have higher socioeconomic status than the general cancer population (78), these consistent biases might enable a more, rather than less, fair comparison between trial and nontrial patients with respect to survival. Unfortunately, socioeconomic status was not available for both databases. Moreover, the nature of SEER data, with respect to racial, ethnic, sociodemographic, and age distributions,

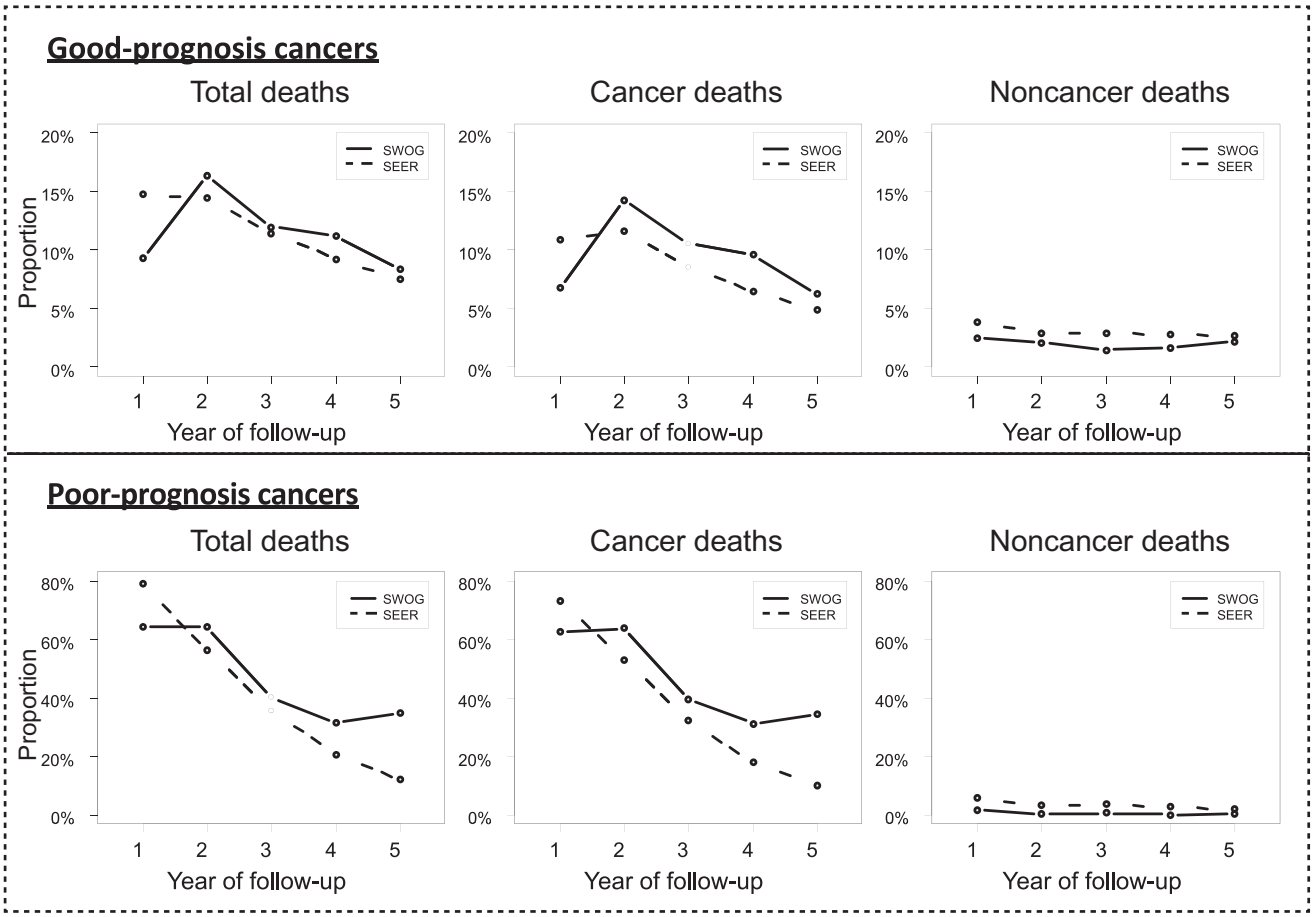

Figure 5. Total, cancer-specific, and non-cancer-specific deaths by year of follow-up by prognosis. For each of the first 5 years, the proportion of patients experiencing death of any kind, cancer-specific death, and non-cancer-specific death relative to the number of patients at risk in each year is plotted for both SWOG and Surveillance, Epidemiology, and End Results (SEER) patients. Consistent with the Kaplan–Meier survival plots in Figure 4, the total event rate is notably lower in SWOG patients in the first year. In years 2 to 5, in contrast, the proportions of total events in SWOG and SEER patients are more similar and are decreasing as the risk of death decreases. For both good- and poor-prognosis patients, the pattern of a relatively lower event rate for SWOG patients in year 1 is mostly reflective of a diminished rate of cancer-related deaths in year 1; although non-cancer-related deaths are also lower in SWOG patients, this difference was small and relatively stable across all 5 years of follow-up. Indeed, in good-prognosis patients, the unweighted ratio of the rate of SEER cancer deaths to SWOG cancer deaths was 1.60 in year 1 but was less than one (only 0.78) in years 2 to 5. For non-cancer-specific deaths, the ratios are very similar whether in year 1 (1.57) or years 2 to 5 (1.52), indicating the pattern change over time occurs in cancer-related deaths only. A similar pattern held for poor-prognosis cancers. In summary, the difference in the patterns of death for trial vs nontrial patients between year 1 vs. years 2 to 5 is largely attributable to different patterns of cancer deaths.

has changed over time, which could impact analyses in unknown ways, although, importantly, we did not observe temporal trends toward greater or lesser generalizability over calendar time. In addition, these results may not apply to other clinical settings (ie, screening). Finally, some elements of this analysis were not prespecified, so a similar analysis in a different set of studies might reveal, in particular, a different duration of trial benefit than the 1-year effect found in this analysis.

This study also had particular strengths compared with prior studies. The approach of systematically examining an entire cooperative group phase III clinical trial database limited potential subjective selection of studies. It also provided a large panel of studies for comparison. Because these studies were from one cooperative group, other potential sources of variation (eg, data collection methods, payment methods, study designs) were implicitly controlled for. These advantages allowed us to aggregate data across studies and thus distinguish the different behaviors of the survival functions between trial and nontrial patients.

These results may serve as a stimulus to design randomized trials with less strict eligibility criteria (79). We found that eligibility pertaining to comorbid conditions comprised approximately 60% of all criteria. Despite this, histology and stage were primarily determinative of survival outcomes, even in the first year when the influence of trial participation was strongest. Eligibility criteria in clinical trials are clearly required to maintain patient safety; however, consideration should be given to relaxing or eliminating criteria where possible. For instance, laboratory cutoff values may exclude patients who are otherwise clinically appropriate for trial treatment, or the exclusion of patients with prior cancer may be less meaningful in an era in which increasingly more patients are cancer survivors. One concern is that broader eligibility will introduce heterogeneity into the clinical trial cohort, which could reduce statistical power.

Table 3. Main results and prognosis for individual studies included in reviews by Edwards et al. (42) and Peppercorn et al. (41)*

Article†	Cancer type	Results‡	Prognosis group§	Evidence‖	
				Any (U or M)	M¶,#
Antman (E) (47)**	Sarcoma	No U result; no SS difference in M DFS (P = .15); OS not reported	Good	No	No
Bertelsen (E) (48)	Ovarian	Difference in OS in U setting (P < .001) but not M setting w/ same TX (P = .98)	Good	Yes	No
Boros (P) (49)	AML	Difference in OS in U setting (P < .001) and in M setting (P = .02)	Poor	Yes	Yes
Burgers (P) (50)	SCLC	No SS difference in OS in U (no P value given); M not done	Poor	No	—
Cottin (P) (51)	SCLC	SS difference in the U (P = .01) but not M setting (unknown P value); adjusted for performance status	Poor	Yes	No
Dahlberg (P) (52)	Rectal	No differences between trial and nontrial pts of similar TX (surgery)	Good	No	No
Davis (B) (53)	NSCLC	SS difference in both U (P < .001) and M setting (P < .002)	Good	Yes	Yes
Dowling (P) (54)	Prostate	SS difference in U (P = .003) but not M setting after adjusting for performance status (P = .42)	Poor	Yes	No
Feuer (P) (55)	1) Testicular	Minimal disease: SS difference in both U and M	Good	Yes	Yes
	2) Testicular	Advanced disease: No difference in U or M	Good	No	No
Greil (P) (56)	Hodgkin's	No difference in OS in either U (P = .67) or M (P = .65) settings	Good	No	No
Karjalainen (B) (57)	1) Myeloma	1979–85: SS difference in favor of trial pts	Good	Yes	Yes
	2) Myeloma	1959–78: NS trend in favor of nontrial pts	Good	No	No
Lennox (B) (58)	Wilms††	SS difference in OS in both U (P < .01) and M settings (P < .001)	Good	Yes	Yes
Link (P) (59)	Osteo-sarcoma††	No difference in OS in U (no P value)	Good	No	—
Marubini (P) (60)	Breast	SS in U setting (no P value given) but not M setting (P = .50)	Good	Yes	No
Mayers (P) (61)	Breast	SS in U setting (P = .02) but not M setting (P = .09)	Good	Yes	No
Meadows (P) (62)	ALL††	SS differences in U (P < .001) and M (no P value) settings	Good	Yes	Yes
MRC (E) (63)	Leukemia††	Difference in OS (P value not given)	Poor	Yes	—
Roy (P) (64)	Hodgkin's	No P values given. OS appears worse for nontrial pts in older (≥45 y) but not younger pts	Good	Yes	—
Schea (P) (65)	SCLC	SS difference in U (P = .002)	Poor	Yes	—
Schmoor (B) (66)	Breast	Trial 2) No difference in DFS in U	Good	No	—
		Trial 3) NS DFS trend in favor of trial pts in U	Good	No	—
Stiller (P) (67)	ALL††	No difference in U (P = .63)	Good	No	—
	AML††	SS difference in U (p=.04); in M, No difference in 1984–1988, Difference in 1989–1994	Poor	Yes	Yes
Stiller (B) (68)	ALL††	SS difference for both U (no P value given) and M (P < .0001)	Good	Yes	Yes
Stiller (B) (69)	AML††	1975–83: U not done; SS difference in M (p<.001)	Poor	Yes	Yes
		1984–88: U not done; No difference in M	Good	No	No
Stiller (P) (70)	ALL††	1980–84: U not done; No difference in M (P = .62)	Good	Yes	No
		1985–89: U not done; Difference in M (P = .02)	Good	Yes	Yes
		1990–94: U not done; Difference in M (P < .0001)	Good	Yes	Yes
Wagner (P) (71)	NHL††	SPOG vs nonstudy: No SS difference in U (P = .07)	Good	No	—
		POG vs nonstudy: SS difference in U (P < .0001)	Good	Yes	—
Ward (B) (72)	Stomach	5/10 analyses were SS (P ≤ .05; Table III)	Poor	Yes	—
Winger (P) (73)	Glioma	SS difference in U (P = .00001) vs all nonstudy pts	Poor	Yes	—
	Glioma	NS for U (P = .12) vs all nonstudy pts	Poor	No	—

* ACM = all-cause mortality; ALL =; AML = acute myeloid leukemia; AS = actuarial survival; DFS = disease-free survival; M = multivariable; NHL = Non-Hodgkin's lymphoma; NS = nonsignificant; NSCLC = non–small cell lung cancer; OS = overall survival; pts = patients; SCLC = small cell lung cancer; SS = statistically significant; TX = treatment; U = univariate.

† "E" indicates article was included in Edwards et al. (42), "P" indicates article was included in Peppercorn et al. (41), and "B" indicates article was included in both reviews.

‡ Results based on overall survival for all studies except Antman et al. (47) and Schmoor et al. (66).

§ Prognosis groups: Good prognosis is defined as 50% or greater average estimated 2-year survival. Poor prognosis is defined as less than 50% average estimated 2-year survival.

‖ Consistent with our own analysis, studies were categorized according to whether there was a statistically significant (P < .05) difference between trial and nontrial patients.

¶ Among studies where multivariable analyses were conducted.

A dash indicates that no multivariable analyses were conducted.

** Based on full published article for the conference abstract cited by both authors.

†† Childhood cancer.

However, because histology and stage are the dominant predictors of outcome, sufficient homogeneity will be retained even if less impactful criteria are softened. Expanding eligibility would have the further advantage of increasing access to clinical trials for a broader cross-section of patients.

References

1. Tejada HA, Green SB, Trimble EL, et al. Representation of African-Americans, Hispanics, and whites in National Cancer Institute cancer treatment trials. *J Natl Cancer Inst*. 1996;88(12):812–816.
2. Murthy VH, Krumholz HM, Gross CP. Participation in cancer clinical trials: race-, sex-, and age-based disparities. *JAMA*. 2004;291(22):2720–2726.
3. Ford JG, Howerton HW, Lai GY, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. *Cancer*. 2008;112(2):228–242.
4. Ross S, Grant A, Counsell C, Gillespie W, Russell I, Prescott R. Barriers to participation in randomised controlled trials: a systematic review. *J Clin Epidemiol*. 1999;52(12):1143–1156.
5. Unger JM, Green S, Albain KS. Under-representation of elderly patients in cancer clinical trials: causes and remedial strategies. In: Balducci L, Lyman GH, Ershler WB, Extermann M, eds. *Comprehensive Geriatric Oncology*. 2nd ed. Taylor and Francis; 2004:464–491.
6. Begg CB, Zelen M, Carbone PP, et al. Cooperative groups and community hospitals. Measurement of impact in the community hospitals. *Cancer*. 1983;52(9):1760–1767.
7. Hunter CP, Frelick RW, Feldman AR, et al. Selection factors in clinical trials: results from the Community Clinical Oncology Program Physician's Patient Log. *Cancer Treat Rep*. 1987;71(6):559–565.
8. Javid SH, Unger JM, Gralow JR, et al. A prospective analysis of the influence of older age on physician and patient decision-making when considering enrollment in breast cancer clinical trials (SWOG S0316). *Oncologist*. 2012;17(9):1180–1190.
9. Klabunde CN, Springer BC, Butler B, White MS, Atkins J. Factors influencing enrollment in clinical trials for cancer treatment. *South* Med J. 1999;92(12):1189–1193.
10. Green S, Benedetti J, Crowley J. *Clinical Trials in Oncology*. 2nd ed. Boca Raton, FL: CRC Press; 2003.
11. Newhouse JP, McClellan M. Econometrics in outcomes research: The use of instrumental variables. *Annu Rev Public Health*. 1998;19:17–34.
12. Ries LAG, Melbert D, Krapcho M, et al., eds. *SEER Cancer Statistics Review, 1975–2005*. Bethesda, MD: National Cancer Institute; 2008.
13. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457–481.
14. Cox D. Regression models and life tables. *J R Stat Assoc*. 1972;34(2):187–220.
15. Gefeller O, Dette H. Nearest neighbor kernel estimation of the hazard function from censored data. *J Statist Comput Simul*. 1992;43(1–2):93–101.
16. Hess KR, Serachitopol DM, Brown BW. Hazard function estimators: a simulation study. *Stat Med*. 1999;18(22):3075–3088.
17. Mueller HG, Wang JL. Hazard rates estimation under random censoring with varying kernels and bandwidths. *Biometrics*. 1994;50(1):61–76.
18. Anderson JE, Kopecky KJ, Willman CL, et al. Outcome after induction chemotherapy for older patients with acute myeloid leukemia is not improved with mitoxantrone and etoposide compared to cytarabine and daunorubicin: a Southwest Oncology Group study. *Blood*. 2002;100(12):3869–3876.
19. Blumenthal DT, Wade M, Rankin C, et al. MGMT methylation in newly-diagnosed glioblastoma multiforme (GBM): from the S0001 phase III study of radiation therapy (RT) and O-benzylguanine (O BG) plus BCNU versus RT and BCNU alone for newly diagnosed GBM. *J Clin Oncol.* 2006;24(18S):1512.
20. Eisenberger MA, Blumenstein BA, Crawford ED, et al. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. *N Eng J Cancer*. 1998;339(15):1036–1042.
21. Ellis GK, Barlow WE, Gralow JR, et al. Phase III comparison of standard doxorubicin and cyclophosphamide versus weekly doxorubicin and daily oral cyclophosphamide plus granulocyte colony-stimulating factor as neoadjuvant therapy for inflammatory and locally advanced breast cancer: SWOG 0012. *J Clin Oncol*. 2011;29(8):1014–1021.
22. Flanigan RC, Salmon SE, Blumenstein BA, et al. Nephrectomy followed by interferon alfa 2-b compared with interferon alfa-2b alone for metastatic renal-cell cancer. *N Eng J Med*. 2001;345(23):1655–1659.
23. Gandara DR, Crowley J, Livingston RB, et al. Evaluation of cisplatin intensity in metastatic non-small cell lung cancer: a phase III study of the Southwest Oncology Group. *J Clin Oncol*. 1993;11(5):873–878.
24. Godwin JE, Kopecky KJ, Head DR, et al. A double-blind placebo-controlled trial of granulocyte colony-stimulating factor in elderly patients with previously untreated acute myeloid leukemia: a Southwest Oncology Group study (9031). *Blood*. 1998;91(10):3607–3615.
25. Grossman HB, Natale RB, Tangen CM, et al. Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer. *N Eng J Med*. 2003;349(9):859–866.
26. Kelly K, Crowley J, Bunn PA, et al. Randomized phase III trial of paclitaxel plus carboplatin versus vinorelbine plus cisplatin in the treatment of patients with advanced non-small cell lung cancer: a Southwest Oncology Group trial. *J Clin Oncol*. 2001;19(13):3210–3218.
27. Lamm DL, Blumenstein BA, Crawford ED, et al. Randomized intergroup comparison of bacillus calmette-guerin immunotherapy and mitomycin C chemotherapy prophylaxis in superficial transitional cell carcinoma of the bladder. A Southwest Oncology Group study. *Urol Oncol*. 1995;1(3):119–126.
28. Lara PN Jr, Natale R, Crowley J, et al. Phase III trial of irinotecan/cisplatin compared with etoposide/cisplatin in extensive-stage small-cell lung cancer: clinical and pharmacogenomic results from SWOG S0124. *J Clin Oncol*. 2009;27(15):2530–2535.
29. Linden HM, Haskell CM, Green S, et al. Sequenced compared with simultaneous anthracycline and cyclophosphamide in high-risk stage I and II breast cancer: final analysis from INT-0137 (S9313). *J Clin Oncol*. 2007;25(6):656–661.
30. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, Haller DG, Ajani JA, Gunderson LL, Jessup JM, Martenson JA. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N Eng J Med*. 2001;345(10):725–730.
31. MeyskensJr FL, Kopecky KJ, Taylor CW, et al. Randomized trial of adjuvant human interferon gamma versus observation in high-risk cutaneous melanoma: a Southwest Oncology Group study. *J Natl Cancer Inst*. 1995;87(22):1710–1713.
32. PetersIII WA, Liu PY, Barrett RJ, et al. Concurrent chemotherapy and pelvic radiation therapy compared with pelvic radiation thereapy alone as adjuvant therapy after radical surgery in high-risk early stage cancer of the cervix. *J Clin Oncol*. 2000;18(8):1606–1613.
33. Philip PA, Benedetti J, Corless CL, 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(22):3605–3610.
34. Pisters KM, Vallières E, Crowley JJ, et al. Surgery with or without preoperative paclitaxel and carboplatin in early-stage non-small-cell lung cancer: Southwest Oncology Group Trial S9900, an intergroup, randomized, phase III trial. *J Clin Oncol*. 2010;28(11):1843–1849.
35. Salmon SE, Crowley JJ, Grogan TM, Finley P, Pugh RP, Barlogie B. Combination chemotherapy, glucocorticoids, and interferon alfa in the treatment of multiple myeloma: a Southwest Oncology Group study. *J Clin Oncol*. 1994;12(11):2405–2414.
36. Sondak VK, Liu PY, Tuthill RJ, et al. Adjuvant immunotherapy of resected, intermediate-thickness, node-negative melanoma with an allogeneic tumor vaccine: overall results of a randomized trial of the Southwest Oncology Group. *J Clin Oncol*. 2002;20(8):2058–2066.
37. Williamson SK, Crowley JJ, Lara Jr PN, et al. Phase III trial of paclitaxel plus carboplatin with or without tirapazamine in advanced non-small-cell lung cancer: Southwest Oncology Group trial S0003. *J Clin Oncol*. 2005;23(36):9097–9104.
38. Wozniak AJ, Crowley JJ, Balcerzak SP, et al. Randomized trial comparing cisplatin with cisplatin plus vinorelbine in the treatment of advanced non-small-cell lung cancer: a Southwest Oncology Group study. *J Clin Oncol*. 1998;16(7):2459–2465.

39. Hutchins LF, Unger JM, Crowley JJ, Coltman CA Jr, Albain KS. Underrepresentation of patients 65 of age or older in cancer-treatment trials. *N Eng J Med*. 1999;341(27):2061–2067.
40. Unger JM, Coltman CA Jr., Crowley JJ, et al. Impact of the year 2000 Medicare policy change on older patient enrollment to cancer clinical trials. *J Clin Oncol*. 2006;24(1):141–144.
41. Peppercorn JM, Weeks JC, Cook EF, Joffe S. Comparison of outcomes in cancer patients treated within and outside clinical trials: conceptual framework and structured review. *Lancet*. 2004;363(9405):263–270.
42. Edwards SJ, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J. Ethical issues in the design and conduct of randomised controlled trials. *Health Technol Assess*. 1998;2(15):1–132.
43. Elting LS, Cooksley C, Bekele BN, et al. Generalizability of cancer clinical trial results: prognostic differences between participants and nonparticipants. *Cancer*. 2006;106(11):2452–2458.
44. Stiller CA. Centralised treatment, entry to trials and survival. *Br J Cancer*. 1994;70(2):352–362.
45. Tanai C, Nakajima TE, Nagashima K, et al. Characteristics and outcomes of patients with advanced gastric cancer who declined to participate in a randomized clinical chemotherapy trial. *J Oncol Pract*. 2011;7(3):148–153.
46. Vist GE, Hagen KB, Devereaux PJ, Bryant D, Kristoffersen DT, Oxman AD. Systematic review to determine whether participation in a trial influences outcome. *BMJ*. 2005;330(7501):1175–1181.
47. Antman K, Amato D, Wood W, et al. Selection bias in clinical trials. *J Clin Oncol*. 1985;3(8):1142–1147.
48. Bertelsen K. Protocol allocation and exclusion in two Danish randomized trials in ovarian cancer. *British J Cancer*. 1991;64:1172–1176.
49. Boros L, Chuang C, Butler FO, Bennett JM. Leukemia in Rochester (NY). A 17-year experience with an analysis of the role of cooperative group (ECOG) participation. *Cancer*. 1985;56(9):2161–2169.
50. Burgers JA, Arance A, Ashcroft L, Hodgetts J, Lomax L, Thatcher N. Identical chemotherapy schedules given on and off trial protocol in small cell lung cancer: response and survival results. *Br J Cancer*. 2002;87(5):562–566.
51. Cottin V, Arpin D, Lasset C, et al. Small-cell lung cancer: patients included in clinical trials are not representative of the patient population as a whole. *Ann Oncol*. 1999;10(7):809–815.
52. Dahlberg M, Glimelius B, Pahlman L. Improved survival and reduction in local failure rates after preoperative radiotherapy: evidence for the generalizability of the results of Swedish Rectal Cancer Trial. *Ann Surg*. 1999;229(4):493–497.
53. Davis S, Wright PW, Schulman SF, et al. Participants in prospective, randomized clinical trails for resected non-small cell lung cancer have improved survival compared with nonparticipants in such trials. *Cancer*. 1985;56(7):1710–1718.
54. Dowling AJ, Czaykowski PM, Krahn MD, Moore MJ, Tannock IF. Prostate specific antigen response to mitoxantrone and prednisone in patients with refractory prostate cancer: prognostic factors and generalizability of a multicenter trial to clinical practice. *J Urol*. 2000;163(5):1481–1485.
55. Feuer EJ, Frey CM, Brawley OW, et al. After a treatment breakthrough: a comparison of trial and population-based data for advacned testicular cancer. *J Clin Oncol*. 1994;12(2):368–377.
56. Greil R, Holzner B, Kemmler G, et al. Retrospective assessment of quality of life and treatment outcome in patients with Hodgkin's disease from 1969 to 1994. *Eur J Cancer*. 1999;35(5):698–706.
57. Karjalainen S, Palva I. Do treatment protocols improve end results? A study of survival of patients with multiple myeloma in Finland. *BMJ*. 1989;299(6707):1069–1072.
58. Lennox EL, Stiller CA, Jones PH, Wilson LM. Nephroblastoma: treatment during 1970–73 and the effect on survival of inclusion in the first MRC trial. *Br Med J*. 1979;2(6190):567–569.
59. Link MP, Goorin AM, Miser AW, et al. The effect of adjuvant chemotherapy on relapse-free survival in patients with osteosarcoma of the extremity. *N Engl J Med*. 1986;314(25):1600–1606.
60. Marubini E, Mariani L, Salvadori B, et al. Results of a breast-cancer-surgery trial compared with observational data from routine practice. *Lancet*. 1996;347(9007):1000–1003.
61. Mayers C, Panzarella T, Tannock IF. Analysis of the prognostic effects of inclusion in a clinical trial and of myelosuppression on survival after adjuvant chemotherapy for breast carcinoma. *Cancer*. 2001;91(12):2246–2257.
62. Meadows AT, Kramer S, Hopsono R, Lustbader E, Jarrett P, Evans AE. Survival in childhood acute lymphocytic leukemia: effect of protocol and place of treatment. *Cancer Invest*. 1983;1(1):49–55.
63. MRC Working Group on Leukaemia. Duration of survival of children with acute leukemia. Report to the Medical Research Council from the Committee on Leukaemia and the Working Party on Leukaemia in Childhood. *BMJ*. 1971;4(5778):7–9.
64. Roy P, Vaughan Hudson G, Vaughan Hudson B, Esteve J, Swerdlow AJ. Long-term survival in Hodgkin's disease patients. A comparison of relative survival in patients in trials and those recorded in population-based cancer registries. *Eur J Cancer*. 2000;36(3):384–389.
65. Schea RA, Perkins P, Allen PK, Komaki R, Cox JD. Limited-stage small-cell lung cancer: patient survival after combined chemotherapy and radiation therapy with and without treatment protocols. *Radiology*. 1995;197(3):859–862.
66. Schmoor C, Olschewski M, Schumacher M. Randomized and non-randomized patients in clinical trials: experiences with comprehensive cohort studies. *Stat Med*. 1996;15(3): 263–271.
67. Stiller CA, Benjamin S, Cartwright RA, et al. Patterns of care and survival for adolescents and young adults with acute leukaemia—a population-based study. *Br J Cancer*. 1999;79(3–4):658–665.
68. Stiller CA, Draper GJ. Treatment centre size, entry to trials, and survival in acute lymphoblastic leukaemia. *Arch Dis Child*. 1989;64(5):657–661.
69. Stiller CA, Eatock EM. Survival from acute non-lymphocytic leukaemia, 1971–88: a population based study. *Arch Dis Child*. 1994;70(3):219–223.
70. Stiller CA, Eatock EM. Patterns of care and survival for children with acute lymphoblastic leukaemia diagnosed between 1980 and 1994. *Arch Dis Child*. 1999;81(3):202–208.
71. Wagner HP, Dingeldein-Bettler I, Berchthold W, et al. Childhood NHL in Switzerland: incidence and survival of 120 study and 42 non-study patients. *Med Pediatr Oncol*. 1995;24(5):281–286.
72. Ward LC, Fielding JW, Dunn JA, Kelly KA. The selection of cases for randomised trials: a registry survey of concurrent trial and non-trial patients. The British Stomach Cancer Group. *Br J Cancer*. 1992;66(5):943–950.
73. Winger MJ, Macdonald DR, Schold SC Jr, Cairncross JG. Selection bias in clinical trials of anaplastic glioma. *Ann Neurol.* 1989;26(4):531–534.
74. Braunholtz DA, Edwards SJ, Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect." *J Clin Epidemiol*. 2001;54(3):217–224.
75. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet*. 1993;342(8883):1317–1322.
76. Nattinger AB, McAuliffe TL, Schapira MM. Generalizability of the surveillance, epidemiology, and end results registry population: factors relevant to epidemiologic and health care research. *J Clin Epidemiol*. 1997;50(8):939–945.
77. Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. *Med Care*. 2002;40(8 Suppl):IV-3–18.
78. Sateren WB, Trimble EL, Abrams J, et al. How sociodemographics, presence of oncology specialists, and hospital cancer programs affect accrual to cancer treatment trials. *J Clin Oncol*. 2002;20(8):2109–2117.
79. George SL. Reducing patient eligibility criteria in cancer clinical trials. *J Clin Oncol*. 1996;14(4):1364–1370.

Funding

This work was supported in part by Public Health Service Cooperative Agreement grants CA32102, CA38926, CA68183, CA20319, and CA37429 awarded by the National Cancer Institute, National Institutes of Health, and Department of Health and Human Services.

Affiliations of authors: SWOG Statistical Center, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA (JMU, WEB, ML); University of Washington, Department of Health Services Research, Seattle, WA (DPM); Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA (SDR, RE); Division of Hematology/Oncology, Columbia University, New York, NY (DLH).