

# The effect of cancer drug vintage on cancer survival and mortality

Economists believe that the development of new products is the main reason why people are better off today than they were several generations ago. In their 1993 book, *Innovation and Growth in the Global Economy*, Grossman and Helpman [1] argued that “innovative goods are better than older products simply because they provide more ‘product services’ in relation to their cost of production.” In their 1996 book, *The Economics of New Goods*, Bresnahan and Gordon [2] stated simply that “new goods are at the heart of economic progress.” In a recent paper, *Measuring the Growth from Better and Better Goods*, Bills [3] makes the case that “much of economic growth occurs through growth in quality as new models of consumer goods replace older, sometimes inferior, models.”

In several papers, Lichtenberg has presented evidence that the introduction and use of new drugs have improved people’s health and increased longevity [4–7]. In the previous report, we also presented results from two studies, estimating the effect on survival of availability of new cancer drugs. In this chapter, we will present more evidence relevant to the hypothesis that treating cancer patients with newer drugs increases cancer survival rates. In other words, the hypothesis is that mean cancer drug vintage has a positive effect on cancer survival rates. The vintage of a drug is defined as the first year in which the drug was launched somewhere in the world. The analysis will be on the basis of the cancer drugs identified for this report.

Figure 1 shows the number of cancer drugs first launched in each year, 1940–2004. Since 1984, the number of drugs launched per decade has been almost six times as high as it was during 1940–1983. Almost half of the 66 drugs were first launched in the United States, Germany, or the UK, see chapter 4.

We will test the hypothesis that treating cancer patients with newer drugs increases cancer survival rates by examining the correlation between drug vintage and survival, controlling for other factors. In principle, the analysis could be performed using individual level data or data on groups of cancer patients. Since individual level data necessary to conduct the analysis were not available, we will perform the analysis using group level data.

We believe that the preferred way to measure mean cancer drug vintage ( $V$ ) with group level data is as follows:

$$V = \sum_d N_d f(\text{YEAR}_d) / \sum_d N_d,$$

where  $N_d$  is the number of patients treated with cancer drug  $d$  ( $d = 1, \dots, 66$ ) and  $f(\text{YEAR}_d)$  is a function of the year of initial launch of cancer drug  $d$ .

One possible function is  $f(\text{YEAR}_d) = \text{YEAR}_d$ . In this case,  $V$  is the weighted-average year of initial launch, weighting by the number of patients treated. Another possible function is

$$f(\text{YEAR}_d) = 1, \text{ if } \text{YEAR}_d \geq 1990, \\ = 0, \text{ if } \text{YEAR}_d < 1990.$$

In this case,  $V$  is the fraction of patients treated with drugs launched after 1990.

Cancer patients, and groups of patients, can be classified by a number of characteristics, including primary cancer site, country (of residence), and year (of diagnosis, treatment, or death). Ideally, we would like to have (three-dimensional) data on mean drug vintage and survival of cancer patients cross-classified by all three attributes, i.e. we would like to know the mean drug vintage and survival of each cancer site in each country in each year. Unfortunately, such comprehensive data are not available. However, we were able to construct three different ‘two-dimensional’ datasets, i.e. datasets in which groups of patients were cross-classified by two characteristics, for a given value of the third characteristic. In particular, we constructed datasets on drug vintage and survival (or mortality):

- by primary cancer site and year, for a given country (the United States),
- by primary cancer site and country, for a given year (2002), and
- by country and year, for all cancer sites combined

These two-dimensional datasets are far superior to one-dimensional datasets (e.g. data by country or data by year) because they enable estimation of difference-in-difference (DD) models, instead of just cross-sectional or time-series models. DD models enable us to control for the influence of potentially confounding variables far better than cross-sectional or time-series models.

In the next three sections, we will describe our analyses of the effect of drug vintage on survival by cancer site and year, cancer site and country, and country and year.

## analysis by primary cancer site and year in the United States

The first analysis will be conducted using data on survival, drug vintage, and other variables, by primary cancer site and year, in the United States during the period 1992–2002. We will use these data to estimate the following econometric model of

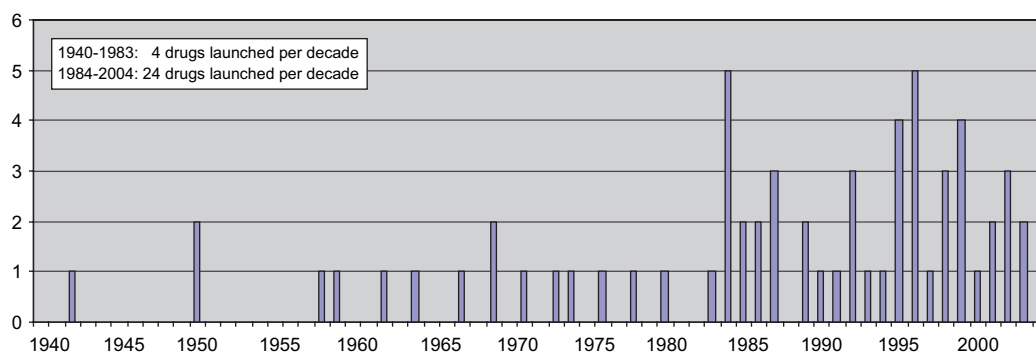


Figure 1. Number of cancer drugs first launched in each year, 1940–2004.

observed survival:

$$f(\text{SURV}_{it}) = \beta \left[ \frac{\sum_d N_{dit} f(\text{YEAR}_d)}{\sum_d N_{dit}} + \gamma Z_{it} + \alpha_i + \delta_t + \epsilon_{it} \right]$$

$$f(\text{SURV}_{it}) = \beta V_{it} + \gamma Z_{it} + \alpha_i + \delta_t + \epsilon_{it} \quad (1)$$

where  $\text{SURV}_{it}$  is the observed survival rate of people diagnosed with cancer originating at site  $i$  in year  $t$ . The observed survival rate is the probability of surviving all causes of death for a specified time interval. Observed survival does not consider cause of death, it simply looks at who is alive and who is not.  $N_{dit}$  is the number of times cancer drug  $d$  was used to treat patients with cancer type  $i$  in year  $t$ ;  $\text{YEAR}_d$  the year of initial launch of cancer drug  $d$ ;  $V_{it} = \frac{\sum_d N_{dit} f(\text{YEAR}_d)}{\sum_d N_{dit}}$ , a measure of the vintage of drugs used to treat cancer type  $i$  in year  $t$ ;  $Z_{it}$  a vector of other attributes of cancer site  $i$  in year  $t$ ;  $\alpha_i$  fixed cancer-site effects; and  $\delta_t$  fixed year effects

We will use this function of launch year:

$$f(\text{YEAR}_d) = 1, \text{ if } \text{YEAR}_d \geq 1990, \\ = 0, \text{ if } \text{YEAR}_d < 1990.$$

Hence  $V_{it}$  is the fraction of patients with cancer type  $i$  treated with drugs launched after 1990 in year  $t$ .

The vector of other attributes of cancer site  $i$  in year  $t$  will include the following variables:

- Percentage of other (noncancer) drug treatments administered in year  $t$  associated with cancer originating at site  $i$  that used drugs approved by the Food and Drug Administration (FDA) after 1990.
- Percentage of radiation oncology procedures performed in year  $t$  associated with cancer originating at site  $i$  whose CPT codes were established by the American Medical Association (AMA) after 1990.
- Percentage of diagnostic radiation procedures performed in year  $t$  associated with cancer originating at site  $i$  whose CPT codes were established by the AMA after 1990.
- Percentage of surgical procedures performed in year  $t$  associated with cancer originating at site  $i$  in year  $t$  whose CPT codes were established by the AMA after 1990.
- The expected survival rate of people diagnosed with cancer originating at site  $i$  in year  $t$ . The expected survival rate is the observed survival rate of a comparable (in terms of race, sex, and age) set of people who do not have cancer.
- The mean age of people diagnosed with cancer originating at site  $i$  in year  $t$ .

- The number of people diagnosed with cancer originating at site  $i$  in year  $t$ .
- The fraction of cancers originating at site  $i$  in year  $t$  that were diagnosed *in situ* (stage 0).
- The fraction of cancers originating at site  $i$  in year  $t$  that were diagnosed as localized or regional (stage 1 or 2).
- The fraction of cancers originating at site  $i$  in year  $t$  that were diagnosed as distant (stage 4).

Due to the presence of fixed cancer-site effects and year effects, this is a DD model. A positive and significant estimate of  $\beta$  would signify that there were above-average increases in observed survival rates of cancer sites with above-average increases in  $\text{drug\_post1990\%}$ , *ceteris paribus*.

### survival data

Data on observed and expected survival rates, the number of people diagnosed, mean age at diagnosis, and stage distribution were obtained from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) 1973–2003 Public Use Data. I used data from SEER 9 registries, which are Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, and Utah. In this dataset, cases diagnosed from 1973 through 2003 are available for all registries except Seattle-Puget Sound (1974+) and Atlanta (1975+). The database contains one record for each of 3 260 176 tumors. However, data on  $N_{dit}$  are available only for the period 1992–2003.

Cancer cases were classified using the SEER Cancer Causes of Death Recode 1969+ (17 September 2004) shown in Appendix Table 1. This classification includes 68 nonoverlapping types of cancer.

Survival rates may be calculated for a variety of time intervals (1 year, 2 year, etc.). We will estimate models of 1-year, 2-year, and 3-year survival rates. The longer the time interval, the shorter is the available time series. Table 1 shows 1992 and 2000 2-year survival data for the top 50 (ranked by 1992 incidence) cancer sites, or groups of cancer sites. The 2-year relative survival rate for all cancer sites combined increased from 72% in 1992 to 75% in 2000. The 1992–2002 change in survival rates varied considerably across cancer sites. For example, the relative survival rate for female genital system remained constant at 83%, while the relative survival rate for skin excluding basal and squamous increased from 83% to 97%. The relative survival rate for

**Table 1.** 1992 and 2000 survival data, top 50 (ranked by 1992 incidence) cancer sites

Site	Number diagnosed		Two-year survival rates (%)				Relative	
	1992	2000	Observed		Expected		1992	2000
			1992	2000	1992	2000		
All sites	102 446	107 989	67	71	93	94	72	75
Male genital system	20 608	18 385	89	92	90	92	99	100
Prostate	19 832	17 449	89	92	90	92	99	100
Digestive system	17 716	18 875	52	52	92	92	56	57
Breast	15 660	19 450	91	93	96	96	95	97
Respiratory system	14 105	13 735	27	28	93	93	29	30
Lung and bronchus	12 710	12 455	23	25	93	93	25	27
Colon and rectum	11 095	11 295	69	70	92	91	75	77
Colon excluding rectum	7935	8020	68	68	91	91	74	75
Female genital system	7115	6723	80	80	96	96	83	83
Urinary system	5972	6703	74	76	92	92	80	83
Skin excluding basal and squamous	5100	6736	80	93	96	95	83	97
Lymphoma	4231	4785	64	70	95	94	67	75
Melanoma of the skin	3901	6199	92	94	96	96	96	99
Urinary bladder	3856	4082	78	79	91	91	86	88
Non-Hodgkin's lymphoma	3559	4095	59	67	94	94	63	72
Rectum and rectosigmoid junction	3160	3275	71	74	93	93	77	80
Sigmoid colon	2809	2459	75	76	92	92	81	82
Corpus and uterus, not otherwise specified	2752	2940	86	86	95	96	90	90
Corpus uteri	2727	2889	87	87	95	96	91	91
Leukemia	2501	2661	54	53	93	93	59	57
NHL—Nodal	2394	2725	59	65	94	94	62	70
Miscellaneous	2300	1991	16	20	91	91	18	22
Oral cavity and pharynx	2296	2251	65	70	94	95	69	74
Rectum	2094	2324	71	74	93	93	77	80
Pancreas	2078	2267	8	10	92	92	9	11
Kidney and renal pelvis	1948	2460	66	70	94	94	70	75
Ovary	1860	1962	64	67	96	96	67	70
Cecum	1753	1839	63	65	90	90	69	73
Cervix uteri	1731	943	85	79	98	98	86	81
Stomach	1693	1690	30	31	92	91	33	34
Brain and other nervous system	1464	1532	38	40	96	96	40	41
Endocrine system	1391	1927	91	94	98	98	93	95
Brain	1376	1410	36	37	96	96	38	38
Thyroid	1263	1779	93	96	98	98	95	97
Lymphocytic leukemia	1253	1209	77	77	93	93	83	82
Other nonepithelial skin	1199	537	42	79	98	94	43	84
NHL—Extranodal	1165	1370	60	70	94	94	64	75
Myeloma	1117	1205	52	55	92	92	57	59
Myeloid and monocytic leukemia	1067	1293	32	34	93	93	35	36
Rectosigmoid junction	1066	951	70	75	93	93	76	80
Ascending colon	1043	1397	67	68	91	90	74	76
Larynx	1011	882	75	74	94	94	80	79
Esophagus	828	1002	23	28	92	92	25	30

leukemia declined from 59% to 57%, while the relative survival rate for non-Hodgkin's lymphoma increased from 63% to 72%.

### drug utilization data

Data on utilization, by drug, diagnosis, and year ( $N_{dit}$ ) were obtained from the MEDSTAT MarketScan database. MEDSTAT contains data on outpatient and inpatient services (procedures) and outpatient prescriptions of hundreds of thousands, or even millions, of individuals.

It is worth distinguishing between two types of drugs: self-administered drugs and drugs administered by physicians and other medical providers (e.g. chemotherapy). Utilization of self-administered drugs is reported in outpatient prescription records (claims). These records generally do not include any information about the patient's diagnosis. In contrast, drugs that are administered by physicians and other medical providers are reported as outpatient and inpatient services (procedures). These records include information about the patient's diagnosis.

**Table 2.** Chemotherapy procedures in 2003, ranked by frequency in MEDSTAT database

Chemotherapy procedure	Number of treatments	Total expenditure (in \$)
J9190-fluorouracil injection	25 578	444 502
J9265-paclitaxel injection	11 587	13 500 000
J9000-doxorubic HCl 10 mg VI chemo	10 445	3 940 000
J9045-carboplatin injection	10 082	11 700 000
J9170-docetaxel	9982	16 400 000
J9355-trastuzumab	9175	11 000 000
J9201-gemcitabine HCl	8118	7 840 000
J9206-irinotecan injection	6071	9 720 000
J9310-rituximab cancer treatment	5275	18 200 000
J9390-vinorelbine tartrate/10 mg	4998	2 260 000
J9093-cyclophosphamide lyophilized	3997	279 332
J9096-cyclophosphamide lyophilized	3373	243 015
J9060-cisplatin 10 mg injection	3340	755 718
J9217-leuprolide acetate suspension	3303	5 100 000
J9182-etoposide 100 mg injection	3288	622 629
J9185-Fludarabine Phosphate injection	3275	1 280 000
J9181-etoposide 10 mg injection	2891	412 888
J9031-bcg live intravesical vaccine	2366	416 056
J9095-cyclophosphamide lyophilized	2363	113 985
J9062-cisplatin 50 mg injection	2279	825 800
J9260-methotrexate sodium injection	2002	47 743
J9370-vincristine sulfate 1 mg injection	1967	128 832
J9350-topotecan	1950	1 910 000
J0207-amifostine	1849	920 271
J9070-cyclophosphamide 100 mg injection	1807	127 223
J9202-goserelin acetate implant	1529	1 480 000
J9360-vinblastine sulfate injection	1416	76 148
J9040-bleomycin sulfate injection	1401	678 763
J9094-cyclophosphamide lyophilized	1218	31 339
J9090-cyclophosphamide 500 mg injection	1166	57 949
J9091-cyclophosphamide 1.0 g injection	1074	68 504
J9250-methotrexate sodium injection	969	22 911
J9001-doxorubicin HCl liposome injection	869	2 290 000
J9208-ifosfomide injection	711	404 349
J9140-dacarbazine 200 mg injection	684	59 426
J9130-dacarbazine 10 mg injection	648	67 398
J9100-cytarabine HCl 100 mg injection	608	50 289
J9080-cyclophosphamide 200 mg injection	461	13 383
J9293-mitoxantrone HCl/5 mg	397	452 246
J9097-cyclophosphamide lyophilized	351	43 123
J9010-alemtuzumab injection	267	414 201
J1190-dexrazoxane HCl injection	195	124 520
J9280-mitomycin 5 mg injection	155	56 017
J9291-mitomycin 40 mg injection	151	124 529
J9200-floxuridine injection	144	71 571
J9020-asparaginase injection	141	14 945
J9320-streptozocin injection	128	26 789
J8521-capecitabine, oral, 500 mg	123	77 665
J9290-mitomycin 20 mg injection	108	46 409
J9110-cytarabine HCl 500 mg injection	105	6739
J9050-carmus bischl nitro injection	98	43 558
J1950-leuprolide acetate/3.75 mg	72	80 282
J9340-thiotepa injection	56	12 569
J9092-cyclophosphamide 2.0 g injection	53	8524

**Table 2.** (Continued)

Chemotherapy procedure	Number of treatments	Total expenditure (in \$)
J9150-daunorubicin	53	43 073
J9230-mechlorethamine HCl injection	50	5608
J8700-temozolmide	47	69 743
J9120-dactinomycin actinomycin D	41	4236
J9065-injection cladribine per 1 mg	40	65 878
J9218-leuprolide acetate injection	38	21 785
J9219-leuprolide acetate implant	34	174 216
J9268-pentostatin injection	34	95 320
J8560-etoposide oral 50 mg	28	11 762
J9211-idarubicin HCl injection	21	36 451
J9160-denileukin diftitox, 300 µg	18	190 957
J9151-daunorubicin citrate liposome	17	10 252
J9300-gemtuzumab ozogamicin	16	96 952
J9245-injection melphalan HCl 50 mg	14	11 345
J9263	11	6782
J8530-cyclophosphamide oral 25 mg	10	536
J8520-capecitabine, oral, 150 mg	8	1096
J9266-pegaspargase/single dose vial	7	11 337
J9395	5	1622
J8610-methotrexate oral 2.5 mg	1	6
J9270-plicamycin (mithramycin) injection	1	1051

For most diseases other than cancer, the vast majority of drugs are self-administered, and determining the diagnosis associated with a particular drug's use can be difficult. But an important fraction of drug treatments for cancer are administered by physicians and other medical providers. We will use data on provider-administered drugs only, since the number of times provider-administered drug  $p$  was used to treat cancer originating at site  $i$  in year  $t$  can be measured precisely.

Each MEDSTAT outpatient and inpatient service record contains one procedure code and one or more ICD-9 diagnosis codes. Codes for drugs administered by providers are Healthcare Common Procedure Coding System (HCPCS) Level II Codes.\* Table 2 shows the top 40 (ranked by frequency) provider-administered drugs associated with all cancer diagnoses in 2003.

We used Multum's Lexicon database (<http://www.multum.com/Lexicon.htm>) to determine the active ingredients of the drugs corresponding to each of these HCPCS Level II Codes. We used data from the Drugs@FDA database (<http://www.fda.gov/cder/drugsatfda/datafiles/default.htm>) to determine the year in which each active ingredient was first approved by the FDA.

\*Level II of the HCPCS is a standardized coding system that is used primarily to identify products, supplies, and services not included in the CPT codes, such as ambulance services and durable medical equipment, prosthetics, orthotics, and supplies when used outside a physician's office. Because Medicare and other insurers cover a variety of services, supplies, and equipment that are not identified by CPT codes, the level II HCPCS codes were established for submitting claims for these items. The development and use of level II of the HCPCS began in the 1980s. Level II Codes are also referred to as alphanumeric codes because they consist of a single alphabetical letter followed by four numeric digits, while CPT codes are identified using five numeric digits. See <http://www.cms.hhs.gov/medicare/hcpcs/codpayproc.asp>.

Table 3 shows the number of post-1990 drug treatments as a percentage of the total number of drug treatments (Post1990%), and the number of provider-administered drugs upon which that statistic is based, for all cancer sites combined during the period 1992–2003.

**Table 3.** Post-1990 share of all drugs administered each year from 1992 to 2003

Year	Post1990%	Number of provider-administered drugs	Number of firms covered by MEDSTAT data
1992	9	17 731	45
1993	12	20 134	45
1994	18	22 516	45
1995	17	29 798	45
1996	18	55 190	92
1997	20	62 235	92
1998	25	101 221	92
1999	29	187 838	95
2000	32	216 476	98
2001	33	230 066	103
2002	32	328 234	200
2003	33	486 409	200

The fraction of post-1990 drugs increased from 9% in 1992 to 33% in 2003. The 1992 figure is on the basis of 17 731 observations (drug treatments), and the 2003 figure is on the basis of 486 409 treatments. The increase in sample size is partly due to the fact that the number of firms covered by the MEDSTAT data increased from 45 in 1992 to 200 in 2003.

## empirical results

We computed weighted least-squares estimates of 12 versions of equation 1: three survival intervals (1-year, 2-year, and 3-year survival), two functional forms (probit and logarithmic), and two sets of weights (the number of MEDSTAT drug treatments and the number of SEER 9 patients diagnosed);  $3 \times 2 \times 2 = 12$ . Table 4 shows estimates of equation 1 for all the three survival intervals, on the basis of the probit functional form and weighted by the number of MEDSTAT drug treatments.

For all the three survival intervals, the coefficient on ‘drug\_new%’ is positive and significant ( $P$  value  $< 0.03$ ). This indicates that the cancer sites whose drug vintage (measured by the share of post-1990 treatments) increased the most during the 1990s tended to have larger increases in observed survival rates, *ceteris paribus*.

As noted above, we also estimated models using an alternative (logarithmic) functional form and an alternative set of weights (the number of SEER 9 patients diagnosed). To

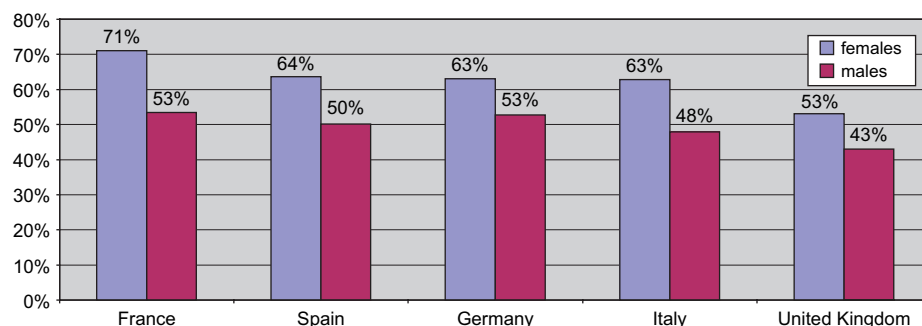
**Table 4.** Estimates of impact of new drugs in model (equation 1)

Parameter	1-year survival			2-year survival			3-year survival		
	Estimate	$t$ value	$P > t$	Estimate	$t$ value	$P > t$	Estimate	$t$ value	$P > t$
drug_new%	0.110	2.62	0.0092	0.145	3.86	0.0001	0.089	2.19	0.0289
rad_diag_new%	0.022	0.18	0.8564	-0.107	-0.93	0.3519	0.054	0.41	0.6839
rad_onc_new%	-0.044	-0.97	0.3343	0.092	2.02	0.0442	0.000	0.01	0.9945
surg_new%	-0.006	-0.04	0.9687	0.003	0.02	0.9834	-0.126	-0.73	0.4678
$F^{-1}(E)$	0.582	4.69	$<0.0001$	0.252	2.24	0.0259	0.410	3.34	0.0009
$\ln(N)$	0.182	5.05	$<0.0001$	0.082	2.18	0.0296	0.205	5.17	$<0.0001$
Age	-0.009	-2.03	0.0425	-0.006	-1.25	0.2131	-0.003	-0.51	0.607
in_situ%	1.240	7.23	$<0.0001$	1.108	6.45	$<0.0001$	1.469	8.30	$<0.0001$
loc_reg%	1.439	11.87	$<0.0001$	0.964	6.32	$<0.0001$	1.381	8.23	$<0.0001$
distant%	1.174	20.83	$<0.0001$	-0.078	-0.38	0.7031	0.280	1.21	0.2276
$\delta_{1992}$	0.030	1.33	0.1845	-0.020	-1.04	0.2974	-0.022	-1.28	0.1998
$\delta_{1993}$	0.015	0.77	0.4402	-0.023	-1.43	0.1534	-0.020	-1.37	0.1708
$\delta_{1994}$	0.019	1.17	0.244	-0.027	-2.02	0.0438	-0.022	-1.89	0.0597
$\delta_{1995}$	0.009	0.57	0.5685	-0.021	-1.57	0.1165	-0.015	-1.42	0.1567
$\delta_{1996}$	0.011	0.83	0.4079	-0.015	-1.35	0.1762	-0.008	-0.93	0.351
$\delta_{1997}$	0.030	2.50	0.0129	-0.003	-0.37	0.7147	0.008	1.03	0.3028
$\delta_{1998}$	0.024	2.42	0.016	-0.004	-0.59	0.5548	0.008	1.65	0.1007
$\delta_{1999}$	0.009	1.15	0.2523	-0.007	-1.26	0.2085	0.000		
$\delta_{2000}$	0.015	2.41	0.0162	0.000					
$\delta_{2001}$	0.000								
Degrees of freedom									
Model	80			79			78		
Error	453			395			338		
Corrected total	533			474			416		
$R^2$	0.999			0.999			0.999		

The dependent variable is  $F^{-1}(S_{it})$ , where  $S_{it}$  is the observed survival rate of people diagnosed with cancer originating at site  $i$  in year  $t$ , and  $F^{-1}(\cdot)$  is the inverse of the standard normal cumulative distribution. All models include cancer-site fixed effects.

**Table 5.** All sites but nonmelanoma skin, males

Country/region	Incidence			Mortality			Prevalence	
	Cases	Crude rate	ASR (W)	Deaths	Crude rate	ASR (W)	1-year	5-year
World	5 801 839	185.7	209.6	3 795 991	121.5	137.7	3 390 545	11 547 465

**Figure 2.** Five-year survival rate, all sites but nonmelanoma skin.

conserve space, we will not present the full estimates of these other nine models. The drug vintage coefficient is positive and significant in 11 of the 12 models and positive and marginally significant ( $P$  value = 0.07) in the other model.

From our estimates, we can calculate the fraction of the 1992–1999 change in the observed survival rate that is attributable to the increased utilization of post-1990 drugs, *ceteris paribus*. The estimated fraction is higher for shorter survival intervals, when observations are weighted by the number of MEDSTAT drug treatments, and logarithmic specification. The mean of the 12 estimates of the fraction of the 1992–1999 change in the observed survival rate that is attributable to the increased utilization of post-1990 drugs is 44%.

## analysis by primary cancer site and country

In this section, we will investigate the effect of cancer drug vintage on survival rates in five countries: France, Germany, Italy, Spain, and UK. We will distinguish between 18 different types of cancer (breast cancer, colon cancer, leukemia, etc.). We will estimate the following model:

$$F^{-1}(\text{SURV}_{ij}) = \beta[\sum_d N_{dij} f(\text{YEAR}_d) / \sum_d N_{dij}] + \alpha_i + \delta_j + \epsilon_{ij}$$

or

$$F^{-1}(\text{SURV}_{ij}) = \beta V_{ij} + \alpha_i + \delta_j + \epsilon_{ij}, \quad (2)$$

where  $\text{SURV}_{ij}$  is the (1-year or 5-year) survival rate for cancer type  $i$  in country  $j$ ;  $F^{-1}(\cdot)$  the inverse of the standard normal cumulative distribution;  $N_{dij}$  the number of times cancer drug  $d$  was used to treat patients with cancer type  $i$  in country  $j$ ;  $\text{YEAR}_d$  the year of initial launch of cancer drug  $d$ ;  $V_{ij} = \sum_d N_{dij} f(\text{YEAR}_d) / \sum_d N_{dij}$ , a measure of the vintage of drugs used to treat cancer type  $i$  in country  $j$ ;  $\alpha_i$  a fixed effect for cancer type  $i$ ;  $\delta_j$  a fixed effect for country  $j$ ; and  $\epsilon_{ij}$  a disturbance.

Note that the model includes both fixed cancer-type effects and fixed country effects. This means that we are controlling for

all determinants of cancer survival that are invariant across cancer types within a given country, and that are invariant across countries for a given cancer type. For example, suppose that rich countries tend to use newer drugs and to have high cancer survival rates, controlling for drug vintage. If we failed to control for country income, we would overstate the impact of drug vintage *per se*. Inclusion of country effects ( $\delta_j$ 's) controls for country income, as well as for other difficult-to-measure country attributes (e.g. the overall quality of cancer treatment) that may influence cancer survival rates and may be correlated with cancer drug vintage.

Due to inclusion of fixed cancer-type and country effects in the model,  $\beta$  represents the effect of 'relative' drug vintage within a country on relative survival rates within the country. Suppose that, on average (across all countries), the survival rate of cancer type A is 25% higher than the survival rate of cancer type B, and the vintage of drugs for cancer type A is 10 years higher than the vintage of drugs for cancer type B. Then one would expect that if, in a particular country, the vintage of drugs for cancer type A is only 5 years higher than the vintage of drugs for cancer type B, that the survival rate of cancer type A is <25% higher than the survival rate of cancer type B. Indeed, estimation of the model requires that the relative vintage of drugs for different cancer types varies across countries.

Due to data limitations, data on survival ( $\text{SURV}_{ij}$ ) and drug utilization ( $N_{dij}$ ) pertain to different time periods. The survival data relate to patients who were diagnosed during 1990–1994 and followed up to 1999. The drug utilization data refer to drugs utilized during 2002Q1–2006Q1. While we would prefer to have survival and drug utilization data for the same period, we hypothesize that survival and drug utilization are reasonably highly serially correlated, e.g. that  $\pi$  is positive and significant in the following equation:

$$\text{VINT}_{2003,ij} = \pi \text{VINT}_{1993,ij} + \eta_i + \tau_j + \kappa_{ij}, \quad (2)$$

where  $\text{VINT}_{2003}$  is the vintage of drugs used in 2003 and  $\text{VINT}_{1993}$  the vintage of drugs used in 1993. In other words,

Table 6.

IMS indication	GLOBOCAN cancer site
Brain	Brain, nervous system
Bronchus and lung	Lung
Chronic lymphocytic leukaemia	Leukaemia
AML	Leukaemia
Colorectal	Colon and rectum
Hodgkins' disease	Hodgkin lymphoma
Kidney, etc.	Kidney
Lung cancer unspecified	Lung
Multiple myeloma and malignant plasma cell	Multiple myeloma
Melanoma of skin	Melanoma
Myeloid leukaemia other/unspecified	Leukaemia
Non-Hodgkin's lymphoma	Non-Hodgkin's lymphoma
Non-small-cell lung cancer	Lung
Other leukaemias	Leukaemia
Ovarian	Ovary, etc.
Small-cell lung cancer	Lung

Table 7.

Country	Number of sample patient drug events
France	17 178
Germany	19 877
Italy	15 099
Spain	13 570
UK	11 537

the vintage of drugs used in 2003 is a noisy but useful indicator of the vintage of drugs used in 1993. In this case, the estimates of the effect of drug vintage on survival are likely to be conservative.

### cancer survival data

The cancer survival data will be drawn from the GLOBOCAN 2002 database. This database provides data on incidence and survival, by country and by cancer site.<sup>†</sup> It has been built up using the huge amount of data available in the Unit of Descriptive Epidemiology of the International Agency for Research on Cancer, part of the World Health Organization. Survival data in GLOBOCAN 2002 were taken directly from the EUROCORE-3 study.<sup>‡</sup> Incidence data are available from cancer registries ([www-dep.iarc.fr/globocan/globocan.html](http://www-dep.iarc.fr/globocan/globocan.html)).

We estimated survival rates by dividing 1-year or 5-year prevalence by incidence (Table 5). One-year survival is

<sup>†</sup>The incidence and survival data are both age adjusted. Generally, the survival data refer to relative survival, which is the probability of not dying from the cancer concerned, excluding other causes of death.

<sup>‡</sup><http://www.eurocare.it/>. The EUROCORE project (European cancer registries study on cancer patients' survival and care) is an international collaborative study on the survival of cancer patients in Europe. It currently involves 67 population-based cancer registries operating in 22 European countries.

Table 8.

Cancer site (recoded)	Number of sample patient drug events	Percentage of sample patient drug events	Cumulative percentage of sample patient drug events
Lung	24 345	31.5	31.5
Breast	18 665	24.2	55.7
Non-Hodgkin's lymphoma	10 305	13.3	69.0
Ovary, etc.	5241	6.8	75.8
Colon and rectum	3328	4.3	80.1
Pancreas	3207	4.2	84.2
Hodgkin lymphoma	2861	3.7	88.0
Bladder	1745	2.3	90.2
Brain, nervous system	1695	2.2	92.4
Multiple myeloma	1626	2.1	94.5
Prostate	1342	1.7	96.2
Stomach	1060	1.4	97.6
Liver	501	0.6	98.3
Oesophagus	490	0.6	98.9
Corpus uteri	371	0.5	99.4
Leukaemia	259	0.3	99.7
Cervix uteri	147	0.2	99.9
Testis	54	0.1	100.0
Thyroid	19	0.0	100.0
Total	77 261	100.0	

estimated to be 58.4% ( $=3390545/5801839$ ), and 5-year survival is estimated to be 39.8% [ $=11547465/(5 \times 5801839)$ ].

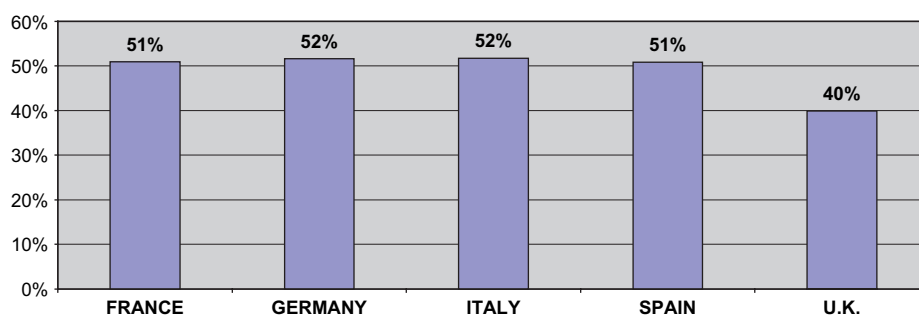
The results of this procedure appear to be consistent with other estimates of survival rates. For example, this implies that the 5-year survival rate for all sites but nonmelanoma skin for males in the United States is 63.8% ( $=2431746/(5 \times 762399)$ ). According to the USA National Cancer Institute, the 5-year survival rate for all sites for males in the United States during 1995–2000 was 64.0%.<sup>§</sup>

Figure 2 shows 5-year survival rates, by country and sex, for all sites but nonmelanoma skin. For both sexes, survival rates in the UK are significantly lower than survival rates in the other four countries.

### drug utilization data

Drug utilization data, by country and cancer site, were obtained from IMS Oncology Analyzer, a diary-derived database of comprehensive patient case history information, on the basis of advanced reporting from hospital clinicians. It provides a full therapeutic itinerary by patient from present day back to diagnosis enabling all aspects of the cancer treatment process to be analyzed, and is on the basis of 70 000 patient cases compiled by >1600 physicians from key specialties across the five European countries and Japan. Coverage of Japan is more recent and limited than that of the other five countries, so we excluded it from the sample to be analyzed. We also recoded some IMS indications (cancer sites) to be consistent with those used in GLOBOCAN, as in Table 6.

<sup>§</sup>Source: [http://seer.cancer.gov/csr/1975\\_2001/results\\_single/sect\\_01\\_table.04\\_2pgs.pdf](http://seer.cancer.gov/csr/1975_2001/results_single/sect_01_table.04_2pgs.pdf).



**Figure 3.** Percentage of 2002–2006 sample patients treated with drugs launched after 1985, by country.

Tables 7 and 8 show the number of sample patient drug events,<sup>¶</sup> by country and by cancer site. Figure 3 shows the percentage of 2002–2006 sample patients treated with drugs launched after 1985, by country. The percentage of post-1985 drugs is significantly lower in the UK than it is in the other four countries.

### estimates

Estimates of equation 2 are shown in Table 9. One-year and 5-year survival models were estimated via weighted least squares, weighting by the number of sample patient drug events  $N_{ij} = \sum_d N_{dij}$ . In the first set of models,  $f(\text{YEAR}_d) = \text{YEAR}_d$ , i.e. the vintage measure is the weighted-average year of initial launch, weighting by the number of sample patient drug events. The vintage variable is not statistically significant in either model.

In the second set of models,  $f(\text{YEAR}_d) = 1$ , if  $\text{YEAR}_d > 1985$ , and  $f(\text{YEAR}_d) = 0$ , if  $\text{YEAR}_d \leq 1985$ , i.e. the vintage measure is the fraction of sample patient drug events in which the drug was launched after 1985. The vintage variable is positive and statistically significant in both the 1-year and 5-year survival models.

The extent to which international differences in survival rates are attributable to differences in drug vintage can be determined by comparing estimates of the country fixed effects ( $\delta_j$ 's) when the vintage variable is included and excluded from the model. This comparison is shown in Figure 4. The difference in the fraction of post-1985 cancer drugs accounts for 14%–19% of the (probit of the) 5-year survival rate differential, adjusted for international differences in distribution of cancer sites. Since the data on survival and drug utilization pertain to different time periods, this estimate is probably conservative.

## analysis by country and year, for all cancer sites combined for 20 countries 1995–2003

Now we will investigate the effect of drug vintage on age-adjusted cancer mortality rates, controlling for several other

<sup>¶</sup>If a sample patient is treated with three drugs, that patient would contribute three patient drug events. We excluded drug events where the cancer site was not explicitly identified (e.g. where the cancer site was designated 'all others').

**Table 9.** Estimates of equation 2

Parameter	Estimate	SD	t value	P
1-year survival, $f(\text{YEAR}_d) = \text{YEAR}_d$				
Drug vintage	-0.017	0.015	-1.14	0.2589
France	0.614	0.058	10.62	<0.0001
Germany	0.358	0.049	7.37	<0.0001
Italy	0.474	0.059	8.10	<0.0001
Spain	0.464	0.052	8.88	<0.0001
UK	0.000			
5-year survival, $f(\text{YEAR}_d) = \text{YEAR}_d$				
Drug vintage	0.002	0.009	0.17	0.8655
France	0.420	0.033	12.86	<0.0001
Germany	0.250	0.029	8.47	<0.0001
Italy	0.315	0.033	9.52	<0.0001
Spain	0.312	0.032	9.85	<0.0001
UK	0.000			
1-year survival, $f(\text{YEAR}_d) = 1$ if $\text{YEAR}_d > 1985$ , $f(\text{YEAR}_d) = 0$ if $\text{YEAR}_d \leq 1985$				
Drug vintage	0.648	0.237	2.74	0.008
France	0.479	0.067	7.18	<0.0001
Germany	0.296	0.051	5.80	<0.0001
Italy	0.354	0.064	5.55	<0.0001
Spain	0.357	0.061	5.84	<0.0001
UK	0.000			
5-year survival, $f(\text{YEAR}_d) = 1$ if $\text{YEAR}_d > 1985$ , $f(\text{YEAR}_d) = 0$ if $\text{YEAR}_d \leq 1985$				
Drug vintage	0.383	0.139	2.76	0.0074
France	0.362	0.036	10.12	<0.0001
Germany	0.216	0.030	7.09	<0.0001
Italy	0.257	0.036	7.05	<0.0001
Spain	0.255	0.036	7.00	<0.0001
UK	0.000			

All models include cancer-site fixed effects.

factors, using annual country level data on 20 countries during the period 1995–2003. We will estimate models of the following form:

$$\ln(M_{ct}) = \beta[\sum_d E_{dct} \text{YEAR}_d / \sum_d E_{dct}] + \gamma Z_{ct} + \alpha_c + \delta_t + \epsilon_{ct}$$

or

$$\ln(M_{ct}) = \beta V_{ct} + \gamma Z_{ct} + \alpha_c + \delta_t + \epsilon_{ct}, \quad (3)$$

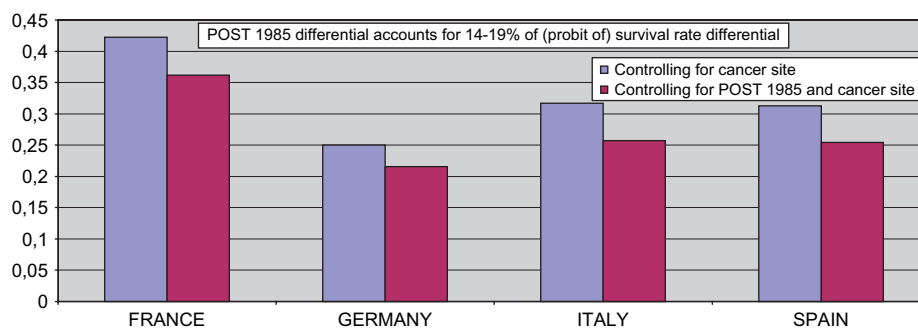


Figure 4. Five-year survival rate differentials from the UK.

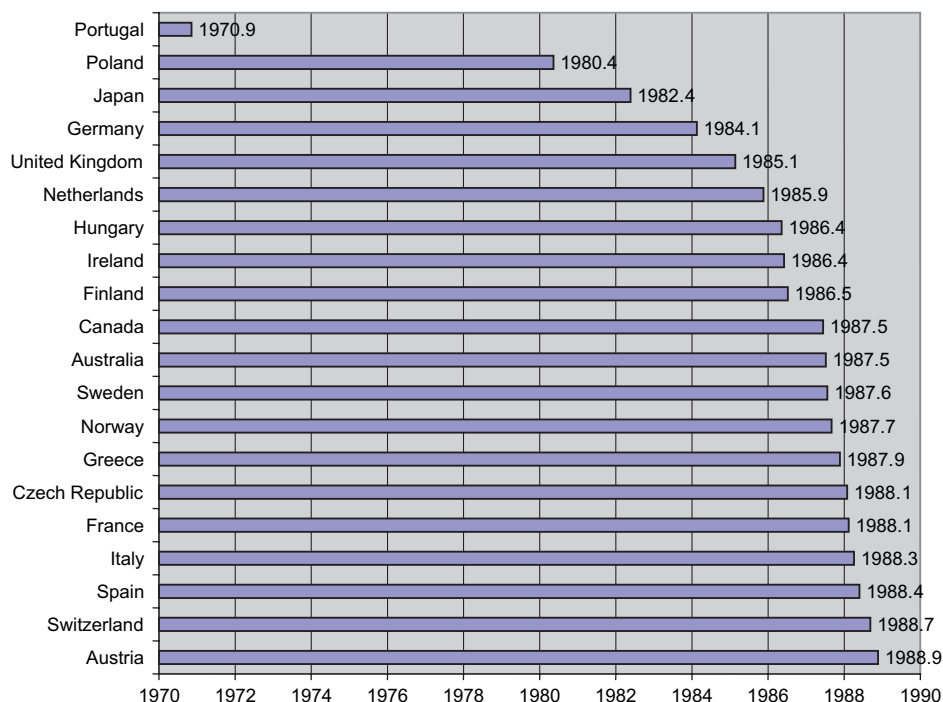


Figure 5. Expenditure-weighted mean launch year of cancer drugs, by country, in 2001.

where  $M_{ct}$  is the age-adjusted mortality rate for all cancer sites combined in country  $c$  in year  $t$ ;  $E_{dct}$  expenditure (in €) on cancer drug  $d$  in country  $c$  in year  $t$ ;  $YEAR_d$  the year of initial launch of cancer drug  $d$ ;  $V_{ct} = \sum_d E_{dct} YEAR_d / \sum_d E_{dct}$ , the weighted-average initial launch year of cancer drugs sold in country  $c$  in year  $t$ , weighted by expenditure;  $Z_{ct}$  other attributes of country  $c$  in year  $t$ ;  $\alpha_c$  a fixed effect for country  $c$ ;  $\delta_t$  a fixed effect for year  $t$ ; and  $\epsilon_{ct}$  a disturbance.

This analysis differs from those in the previous two sections in two important respects. First, our outcome measure is an age-adjusted mortality rate rather than a survival rate. We believe that the survival rate is a superior indicator, but longitudinal country level data on cancer survival rates are not available during the period when drug utilization data are available (1995–2005).<sup>||</sup> Mortality rates depend on incidence as

well as survival. We will therefore estimate a mortality model that controls for (lagged) incidence (by including it in the vector  $Z$ ).

Second, in the vintage measure, drugs are weighted by expenditure (quantity  $\times$  price) rather than by utilization (quantity). We were able to obtain data from IMS Health on expenditure (in €), but not on the number of patients treated, by drug, country, and year. Figure 5 shows the ranking of countries, by expenditure-weighted mean launch year of cancer drugs in 2001.

The fixed country effects control for all determinants of cancer mortality that vary across countries, but not over time. The fixed-year effects control for all determinants of cancer mortality that change over time, but do not vary across countries. It is also desirable to control for other country-specific, time-varying factors that may be correlated with cancer drug vintage. We will control for three such factors.

The first is the age-adjusted cancer incidence rate 8 years earlier. This lag length may be justified on two grounds. One is that incidence data were not available for years after 1997, so

<sup>||</sup>Some longitudinal country level data on cancer survival rates are available for an earlier period from Eurocare, which provides time trends in cancer survival up to 1999 for patients diagnosed during 1983–1994.

choosing an 8-year lag allowed us to maximize the number of observations. The other is that a lag of 8–10 years from incidence to mortality is appropriate; in a previous paper [8], it was estimated that, in the United States, life expectancy at time of cancer diagnosis is ~10 years.

The second covariate is per capita GDP (in PPP USA dollars). It is plausible that countries with larger increases in cancer drug vintage also had higher growth in per capita GDP. If that is the case, and we do not control for per capita GDP,  $\beta$  will capture the effect of higher income as well as newer drugs. This will not necessarily result in overestimation of the effect of vintage on mortality, however. Although faster income growth is likely to be associated with *faster overall* mortality decline, it may well be associated with *slower cancer* mortality decline, due to *competing risks*. Countries with faster income growth may experience substantial reductions in noncancer (e.g. cardiovascular and infectious disease) mortality that are partially offset by a rise in cancer mortality.

The third covariate we will control for is per capita health expenditure (in PPP USA dollars).

Estimates of equation 3 are presented in Table 10. Models were estimated via weighted least squares, where the weight was the population of country  $c$  in year  $t$ .

Model 1 includes no covariates, i.e. the vector  $Z$  is excluded from the model. As expected, the drug vintage coefficient is negative and significant: countries with larger increases in the mean launch year of cancer drugs had larger declines in the age-adjusted cancer mortality rate. The point estimate of  $\beta$  implies that a 10-year increase in drug vintage reduces the mortality rate by 2.5%.

Model 2 includes lagged incidence, per capita health expenditure, and per capita GDP as covariates. The first two variables are far from significant. The coefficient on per capita GDP is positive and highly significant: countries with higher per capita GDP growth had smaller declines in the cancer mortality rate. Controlling for per capita GDP more than doubles the magnitude of the drug coefficient. Model 2 implies that a 10-year increase in drug vintage reduces the mortality rate by 5.9%, controlling for income growth.

From the estimates of Model 2, we can calculate the contribution of the increase in cancer drug vintage to the decline in the age-adjusted cancer mortality rate in the 20 sample countries during the period 1995–2003. The results of this calculation are shown in Figure 6. Controlling for the

effect of per capita GDP growth, the age-adjusted cancer mortality rate declined by ~16% (2% per year) during this 8-year period. The estimates imply that, in the absence of any increase in cancer drug vintage, the age-adjusted cancer mortality rate would have declined by only 11%. Hence the increase in cancer drug vintage—in other words, the use of newer cancer drugs—accounts for ~30% of the GDP growth-adjusted decline in the age-adjusted cancer mortality rate.

Table 10. Estimates of equation 3

Parameter	Estimate	SD	$t$ value	$P$
Model 1: no covariates				
Mean drug launch year (expenditure weighted)	-0.0025	0.001	-2.20	0.029
1995	0.0645	0.016	3.95	0.000
1996	0.0578	0.015	3.77	0.000
1997	0.0431	0.014	3.00	0.003
1998	0.0418	0.014	3.06	0.003
1999	0.0296	0.013	2.27	0.025
2000	0.0189	0.013	1.50	0.136
2001	0.0092	0.012	0.75	0.452
2002	0.0017	0.012	0.14	0.889
2003	0.0000			
Model 2: covariates				
Mean drug launch year (expenditure weighted)	-0.0059	0.001	-4.62	<0.0001
Log incidence rate 8 years earlier	-0.0085	0.050	-0.17	0.866
Log per capita health expenditure (PPP USA dollars)	0.0224	0.057	0.39	0.696
Log per capita GDP (PPP USA dollars)	0.2767	0.052	5.30	<0.0001
1995	0.1206	0.030	4.08	0.000
1996	0.1069	0.027	3.92	0.000
1997	0.0866	0.026	3.35	0.001
1998	0.0822	0.024	3.40	0.001
1999	0.0633	0.023	2.79	0.007
2000	0.0390	0.021	1.86	0.067
2001	0.0214	0.020	1.09	0.280
2002	0.0039	0.019	0.21	0.837
2003	0.0000			

Both models include country fixed effects.

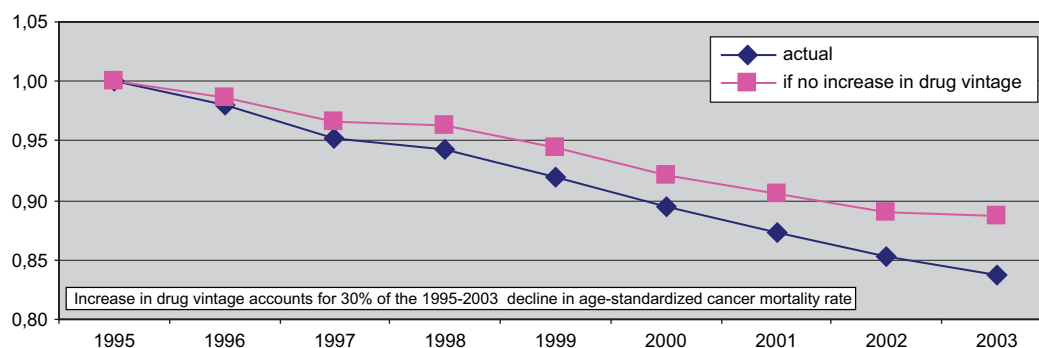


Figure 6. Contribution of the increase in cancer drug vintage to the decline in the age-adjusted cancer mortality rate.

## summary

In the 2005 comparator report, we provided a summary of evidence about the relation between access to new drugs and improvements in survival. Two types of evidence were considered. One study investigated the relation between availability of new drugs and survival for the period 1975–1995 in the United States. A second study compared survival and availability of new drugs in different countries at a defined point in time.

Both studies indicated a positive impact of new drugs. However, the studies were not on the basis of actual data on patient access, measured as utilization or sales of the new drugs. In addition, there was no analysis of the specific impact of the set of defined cancer drugs as identified in this report.

In this chapter, we present three types of evidence about the effect of cancer drug vintage on cancer survival and mortality using actual utilization data. The studies cover different time periods and different countries compared with the studies referred to in the previous report.

All three are on the basis of group level data and employ DD research designs, which enable us to control for the influence of potentially confounding variables far better than cross-sectional or time-series models.

The first and third analyses examine time-series variation in cancer survival/mortality. They address the question: how much of the increase in cancer survival (or decline in cancer mortality) can be attributed to the increasing use of new cancer drugs? The second analysis examines cross-sectional variation in cancer survival. It addresses the question: how much of the international variation in survival of patients in a given period of time can be attributed to international variation in cancer drug vintage?

There is no reason to expect the fraction of time-series variation explained by use of new cancer drugs to be equal to the fraction of cross-sectional variation explained, even in principle. For example, if patients had the same access to new cancer drugs everywhere, then access to new cancer drugs would explain none of the cross-sectional variation in survival, but could still explain a significant part of the time-series variation. Moreover, the three analyses cover different countries and time periods and are on the basis of different measures (dictated by data availability).

The first analysis uses data on cancer drug vintage, survival, and other variables, by primary cancer site and year, for USA cancer patients during the period 1992–2002. We find that the cancer sites whose drug vintage (measured by the share of post-1990 treatments) increased the most during the 1990s tended to have larger increases in observed survival rates, *ceteris paribus*.

The mean estimate of the fraction of the 1992–1999 change in the observed survival rate that is attributable to the increased utilization of post-1990 drugs is 44%.

The second analysis uses data by primary cancer site and country for five large European countries. Drug vintage (in this case measured by the share of post-1985 treatments) has a positive and statistically significant effect on both 1-year and 5-year survival rates. The difference in the fraction of post-1985 cancer drugs accounted for 14%–19% of the 5-year survival rate differential, adjusted for international differences in distribution of cancer sites. Since the data on survival and drug utilization pertain to different time periods, this estimate is probably conservative.

The third analysis is on the basis of data by country and year, for all cancer sites combined, for 20 countries during the period 1995–2003. We find that countries with larger increases in the mean launch year of cancer drugs had larger declines in the age-adjusted cancer mortality rate. The increase in cancer drug vintage—in other words, the use of newer cancer drugs—accounts for ~30% of the GDP growth-adjusted decline in the age-adjusted cancer mortality rate.

## disclosures

Drs Jönsson and Wilking have reported that the publication of this supplement was sponsored by an unrestricted educational grant from Roche Pharmaceuticals.

## references

- Grossman GM, Helpman E. Innovation and Growth in the Global Economy. Cambridge: MIT Press 1993.
- Bresnahan TF, Gordon RJ. The Economics of New Goods. Chicago, IL: University of Chicago Press 1996.
- Bils M. Measuring the growth from better and better goods. NBER Working Paper No. 10606. <http://www.nber.org/papers/w10606>.
- Lichtenberg F. Pharmaceutical knowledge-capital accumulation and longevity. In Corrado C, Haltiwanger J, Sichel D (eds): Measuring Capital in the New Economy. University of Chicago Press 2005; 237–269.
- Lichtenberg F. Availability of new drugs and Americans' ability to work. J Occup Environ Med 2005; 47: 373–380.
- Lichtenberg F. The impact of new drug launches on longevity: evidence from longitudinal disease-level data from 52 countries, 1982–2001. Int J Health Care Finance Econ 2005; 5: 47–73.
- Lichtenberg F. The impact of increased utilization of HIV drugs on longevity and medical expenditure: an assessment based on aggregate U.S. time-series data. Exp Rev Pharmacoecon Outcomes Res 2006; 6: 425–436.
- Lichtenberg F. The Expanding Pharmaceutical Arsenal in the War on Cancer, NBER Working Paper No. 10328, February 2004, <http://www.nber.org/papers/w10328>.