The Local Influence of Pioneer Investigators on Technology Adoption: Evidence from New Cancer Drugs

Leila Agha and David Molitor

February 24, 2017

Appendices

A Appendix: Robustness to continuation mechanism

In the main text of the paper, we interpret our baseline results by arguing that knowledge spillovers from proximity to lead investigators increased new drug utilization in author HRRs. A competing explanation that we discuss in the robustness section of the main paper is that heightened levels of utilization among patients in author HRRs might simply reflect the continuation of treatment for patients enrolled in the trial itself.

In the paper, we discuss results that restrict to new episodes of cancer treatment in an effort to exclude continuing patients from the sample. There are also several other reasons why it is unlikely that the continuation mechanism drives the results we observe. First, while the largest trial sites tend to be located in an author HRR, they are not systematically located in the lead author's HRR.[1] Second, even for a large site with many patients enrolled, it is likely that few of the patients assigned to receive the study drug would show up in our Medicare analysis sample.

To give a back of the envelope sense of the magnitudes that could be explained by continuation effects, we approximate the number of potentially continuing patients in our sample, using information from FDA medical reviews. As a plausible upper bound, we allow the lead author trial sites to enroll 33 patients on average, the average size of the largest site in our analysis of FDA disclosures.[2] Of these 33 patients, assume 16.5 patients are randomized to the treatment arm receiving

[1] We examined FDA medical reviews and clinical trial publications for each of the 21 drugs in our sample to determine, when possible, the total number of patients enrolled in the pivotal clinical trial, the number of trial centers enrolling patients in the trial, the number of patients enrolled in author regions, and the age distribution and median survival of patients in the trial. On average, the pivotal trials for the drugs in our sample enrolled 299 patients across 56 trial sites. While neither the FDA medical reviews nor clinical trial publications systematically report on enrollment by site, select information about trial sites with high enrollment was available for 12 (57%) of the drugs in our study. The average trial size for these twelve drugs was 300 patients, nearly identical to the average trial size of 297 patients for the remaining nine drugs. For the 11 drugs for which we could determine the author location of the largest trial site in the study, 8 (73%) were not located in the lead author's HRR.

[2] Except where otherwise stated, the parameters for this calculation derive from the FDA medical reviews described

the study drug, 8.25 (50%) of these patients survive until after FDA approval, and 2.9 (36%) of the survivors are eligible for Medicare. Even if all 2.9 of these patients continued with the new drug following FDA approval and were counted in our analysis, this would only account for 0.7% of the average base population of 388 indicated cancer patients per drug in the lead author region (Table 2, Column 1, Row 4), only a small fraction of the 4 percentage point increase in new drug utilization we observe in lead author HRRs in the first two years following FDA approval. This calculation further suggests that the continuation mechanism is not likely to explain a significant share of the increase in utilization we estimate in lead author regions.

in the previous footnote. Across the 12 drugs for which site enrollment information was available, the trial site enrolling the most patients enrolled 33 patients, on average. The average age of enrolled patients was approximately 58 years; only 6 of the trials reported the fraction of elderly patients, and in all cases fewer than 36% of patients were 65 years or older. We therefore take 36% as our estimated fraction of trial patients eligible for Medicare. For the 13 drugs reporting median survival in the trial publication, median overall survival in the most favorable treatment arm reported was 14.3 months. This suggests that fewer than half (50%) of individuals enrolled in the pivotal study would have survived to the years following FDA approval of the drug, the baseline years of our analysis.

B Additional Tables and Figures

Table A1: List of Studied Cancer Drugs, Extended Characteristics

Trade name	Generic name	FDA approval date	Journal of publication	Size of target population
(1)	(2)	(3)	(4)	(5)
Xeloda	Capecitabine	4/30/1998	Journal of Clinical Oncology	26,410
Herceptin	Trastuzumab	9/25/1998	Journal of Clinical Oncology	26,410
Valstar	Valrubicin	9/25/1998	Journal of Urology	13,557
Ontak	Denileukin diftitox	2/5/1999	Journal of Clinical Oncology	819
Temodar	Temozolomide	8/11/1999	British Journal of Cancer	1,797
Ellence	Epirubicin hydrochloride	9/15/1999	Journal of Clinical Oncology	53,762
Mylotarg	Gemtuzumab ozogamicin	5/17/2000	Journal of Clinical Oncology	2,192
Trisenox	Arsenic trioxide	9/25/2000	Journal of Clinical Oncology	1,079
Campath	Alemtuzumab	5/7/2001	Blood	12,027
Zometa	Zoledronic acid	8/20/2001	Journal of Clinical Oncology	2,694
Zevalin[1]	Ibritumomab tiuxetan	2/19/2002	Journal of Clinical Oncology	51,042
Faslodex	Fulvestrant	4/25/2002	Journal of Clinical Oncology	64,045
Eloxatin	Oxaliplatin	8/9/2002	Journal of Clinical Oncology	52,778
Velcade[2]	Bortezomib	5/13/2003	New England Journal of Medicine	23,819
Bexxar	Tositumomab-I 131	6/27/2003	Journal of Clinical Oncology	54,275
Alimta	Pemetrexed	2/4/2004	Journal of Clinical Oncology	84,918
Erbitux	Cetuximab	2/12/2004	New England Journal of Medicine	55,528
Avastin	Bevacizumab	2/26/2004	New England Journal of Medicine	55,528
Dacogen	Decitabine	5/2/2006	Cancer	15,460
Arranon	Panitumumab	9/27/2006	Journal of Clinical Oncology	59,028
Torisel	Temsirolimus	5/30/2007	New England Journal of Medicine	3,794

Notes: This table extends Table 1 from the main text, providing additional characteristics on sample drugs.

Table A2: Author Proximity Effect on Drug Utilization: Finer divisions of author role

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug d by patient i

	Panel A: All HRRs		Panel B: Author HRRs only		Panel C: New Cancer Patients	
	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0406*** **		**0.0385*** **		**0.0403*** **	
	(0.0131)		(0.0122)		(0.0154)	
Middle author HRR	**0.0077**		**0.0074**		**0.0076**	
	(0.0050)		(0.0054)		(0.0055)	
Last author HRR	**0.0006**		**0.0028**		**-0.0064**	
	(0.0140)		(0.0127)		(0.0175)	
First author physician group		**0.0422*** **		**0.0407*** **		**0.0422*** **
		(0.0125)		(0.0126)		(0.0162)
First author HRR, non-author group		**0.0417** **		**0.0395** **		**0.0411** **
		(0.0211)		(0.0196)		(0.0203)
Middle author physician group		**0.0273*** **		**0.0273*** **		**0.0259** **
		(0.0074)		(0.0074)		(0.0082)
Middle author HRR, non-author group		**-0.0018**		**-0.0023**		**-0.0007**
		(0.0059)		(0.0064)		(0.0060)
Last author physician group		**0.0324**		**0.0314**		**0.0438**
		(0.0324)		(0.0319)		(0.0329)
Last author HRR, non-author group		**-0.0137**		**-0.0101**		**-0.0275*** **
		(0.0016)		(0.0097)		(0.0087)
Number of observations	659,468	659,468	281,253	281,253	393,618	393,618

Notes: See notes to Table 3. In contrast to results reported in Table 3, we split the "other" author category to separately consider middle and last authors.

Notes on Appendix Table A2

This table reports results from 6 separate regressions that mirror our baseline specification 1, but further subdivide other authors into middle authors and last authors on the basis of authorship order on the academic publication. For new drug clinical trials, the first author is typically the principal investigator and the last author is often a scientist employed by the sponsoring drug company. In fact, only seven drugs in our sample have a last author who is a practicing clinician, whereas all 18 drugs with US-based trials have a practicing clinician as the first author. Note that we only investigate the role of practicing physicians; for the many drugs with a non-clinical final author, there is no last author region coded.

Results reported here find that patients treated in last author regions are not significantly more likely to receive the new drugs compared to other regions; however, the estimates are relatively imprecise due to the small number of drugs with clinicians in the last author position. In particular, the 95% confidence interval includes an up to 2.8 percentage point higher utilization in last author regions, reported from the column (1) baseline specification. Point estimates are suggestive of 3.2 percentage point higher use within the last author's physician group, as reported in column (2), but the findings are not statistically significant.

Table A3: P-values from Alternate Approaches to Inference

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug d by patient i

Proximity Measures	Panel A: Standard cluster at HRR-drug		Panel B: Standard cluster at HRR		Panel C: Wild cluster bootstrap at HRR-drug		Panel D: Wild cluster bootstrap at HRR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First author HRR	0.0404***		0.0404***		0.0404***		0.0404**	
	(0.002)		(0.010)		(0.010)		(0.012)	
Other author HRR	0.0069		0.0069		0.0069		0.0069	
	(0.149)		(0.155)		(0.212)		(0.180)	
First author HRR & in author group		0.0421***		0.0421***		0.0421***		0.0421***
		(0.001)		(0.000)		(0.002)		(0.002)
First author HRR & non-author group		0.0416**		0.0416*		0.0416		0.0416
		(0.048)		(0.076)		(0.168)		(0.278)
Other author HRR & in author group		0.0276***		0.0276***		0.0276***		0.0276***
		(0.000)		(0.002)		(0.002)		(0.002)
Other author HRR & non-author group		-0.0031		-0.0031		-0.0031		-0.0031
		(0.564)		(0.549)		(0.652)		(0.626)
Number of observations	659,468	659,468	659,468	659,468	659,468	659,468	659,468	659,468

Notes: This table explores robustness of our main findings to alternative methods of statistical inference to account for clustering. In columns 1 and 2, we replicate the analysis from Table 3 Panel A, but we now report p-values associated with each coefficient in parentheses. Each subsequent panel reports the same regression results, but applies a different methodology for inference. Panel A uses the usual cluster robust standard error with clusters defined at the hospital referral region (HRR) by drug level. Panel B reports p-values using the usual cluster robust standard error with clusters defined at the HRR. Panel C reports p-values from a wild cluster bootstrap, accounting for clusters at the HRR by drug level. Panel D reports p-values from a wild cluster bootstrap, accounting for clusters at the HRR level.

Table A4: Scope of Author Influence

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug d by patient i

	Regional technology intensity		Neighbor region drug adoption		Off-label drug use	
Independent variables:	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0409*** **	**0.0385*** **	**0.0402*** **	**0.0402*** **	**0.0006**	**0.0007**
	(0.0112)	(0.0105)	(0.0131)	(0.0130)	(0.0005)	(0.0005)
First author HRR * Fast adoption index	**-0.0215** **	**-0.0197*** **				
	(0.0084)	(0.0077)				
Neighbor of first author HRR			**0.0017**	**0.0014**		
			(0.0066)	(0.0066)		
Other author HRR	**0.0064**	**0.0063**	**0.0066**	**0.0067**	**0.0002**	**0.0003**
	(0.0047)	(0.0050)	(0.0048)	(0.0049)	(0.0004)	(0.0004)
Other author HRR * Fast adoption index	**-0.0042**	**-0.0035**				
	(0.0052)	(0.0053)				
Neighbor of other author HRR			**-0.0031**	**-0.0030**		
			(0.0030)	(0.0031)		
Sample						
Restricted sample?	No	Yes	No	Yes	No	Yes
Number of observations	659,468	286,637	659,468	547,256	7,712,248	3,063,237

Notes: This table reports coefficients and standard errors from 6 separate regressions. See notes to Table 3 for further details. Columns 2 and 6 restrict the sample to only include regions that contain at least one study author. Column 4 restricts the sample to include only regions that contain a study author or border an author region. p<0.10; **: p<0.05; ***: p<0.01.

(1-2) Fast adoption index summarizes regional utilization rates of new drugs when there is no author present in the region; the variable is normed to be mean zero and have a standard deviation of 1. Higher values correspond to regions that are quicker to adopt new cancer drugs.

(3-4) Neighbor regions are defined as those that share a border with author regions.

(5-6) Sample is changed to include all cancer care episodes treating patients who do not have the indicated disease type for the observed drug (e.g. testing use of a colon cancer drug in patients with other cancer types, such as breast cancer or lung cancer).

Notes on Appendix Table A4

This table probes the extent of drug authors' influence. In particular, we test which types of regions are most heavily influenced by investigator proximity; whether study authors affect utilization in neighboring regions; and lastly, whether study authors affect off-label drug use.

First, we test whether regions that are typically slow to adopt new cancer drugs experience a greater boost in utilization when a study author is located in the region. There may be greater scope for the study author to affect practice patterns in slower-adopting regions that are not already very high users of new cancer drugs.

We develop a measure of each region's speed of new cancer drug adoption by looking at the average rates of new drug use when no author is present in the region. In particular, we regress an indicator for new drug use on a series of region dummy variables, controlling for drug by year fixed effects, patient demographic characteristics and whether this is a new cancer spell. To avoid

including the direct impact of author proximity in this measure of regional adoption speed, we exclude from the sample any observations where an author for the relevant drug was located in that region. This regression includes only observations in the first two years following initial FDA approval. The region fixed effects from this regression form the basis of our measure of regional technology adoption speed.

For ease of interpretation, we standardize this measure of technology adoption speed by de-meaning the variable and dividing each fixed effect by the standard deviation. As a result, the average regional technology adoption speed index is 0, and a value of 1 corresponds to a region with average new drug use 1 standard deviation above the national mean.

For estimation, we augment our baseline estimating Equation (1) to include interaction terms between whether the first (or other) author is located in the region and the region's technology adoption speed index. Results are reported in Table A4, columns (1-2). We find that the first author's influence is greatest in regions that are typically slower to adopt new drugs; the interaction term is negative and statistically significant at the 5% level. For regions near the mean of the adoption speed index, patients are 4.1 percentage points more likely to receive treatment with the new drug when the first author is located in that region. The impact of being treated in a first author HRR increases to 6.2 percentage points for regions that are typically 1 standard deviation slower to adopt than the average region; the effect falls to 1.9 percentage points for regions that are typically 1 standard deviation faster to adopt.

We continue to find no significant effect of being treated in a middle or last author's region. The coefficient on the interaction between other author HRR and regional adoption speed index similarly suggests that slower-adopting regions experience a greater effect of proximity to other study authors; however, the result is not statistically significant.

The first author's influence boosts regional use more for regions that tend to be technology adoption laggards. From a policy perspective, this suggests that investment in clinical research may yield the greatest spillovers to medical practice in regions that are not already among the fastest adopters of new technologies. It should be noted that this effect is estimated only using variation in adoption speed within the set of regions that contain a first author for at least one in-sample drug.

Next, we test the geographic extent of investigator influence. This analysis not only addresses the geographic extent of the authors' reach, but also impacts the interpretation of our baseline estimates from specification 1. There, we estimated the wedge between investigator HRRs relative

to non-investigator HRRs. If proximity effects extend more broadly than an investigator's own HRR, some of the comparison non-investigator regions are themselves influenced by the treatment, resulting in estimated proximity effects that are too small.

To measure whether investigator influence extends beyond his own HRR, for each drug we identify the "neighbor" HRRs that share a border with the HRR in which the drug's first author is located. We augment our baseline estimating Equation 1 to include two indicator variables: one for patients treated in a region that neighbors a first author region, and a second for patients treated in a region that neighbors a middle or last author region.

Table A4 shows in columns (3-4) that while first author HRRs have a 4.0 percentage point increase in their propensity to use the new drug, there is no observed increase in drug use of neighboring HRRs. The point estimate suggests a less than 0.2 percentage point increase in new drug utilization in neighboring HRRs, which is small in magnitude and statistically not distinguishable from zero. This null effect is quite precise, with a 95% confidence interval that excludes effect sizes that are one third as large as the impact of being treated in the first author's HRR. There is a similarly small, insignificant effect estimated for neighbors of other author HRRs. Although the first author's influence may extend beyond physicians in his own practice group to other physicians practicing in the same region, there is no evidence that his influence raises utilization in neighboring regions.

Finally, we investigate whether study authors influence the use of new drugs for applications not covered by the initial FDA approval label. While drug labels typically provide relatively narrow indications for application, physicians have wide latitude in determining how they will prescribe the drug.[3] Across our 21 drugs, 22% of utilization within the first two years was for patients with diseases not indicated on the FDA label, which we will call "far" off-label drug use. In columns (5-6) of Table A4, we estimate whether the study authors' influence increases the use of new drug for other applications. The sample is restricted to patients who do *not* have the broad cancer type covered by the initial FDA label, and the estimating equation mirrors the specification in Equation 1.

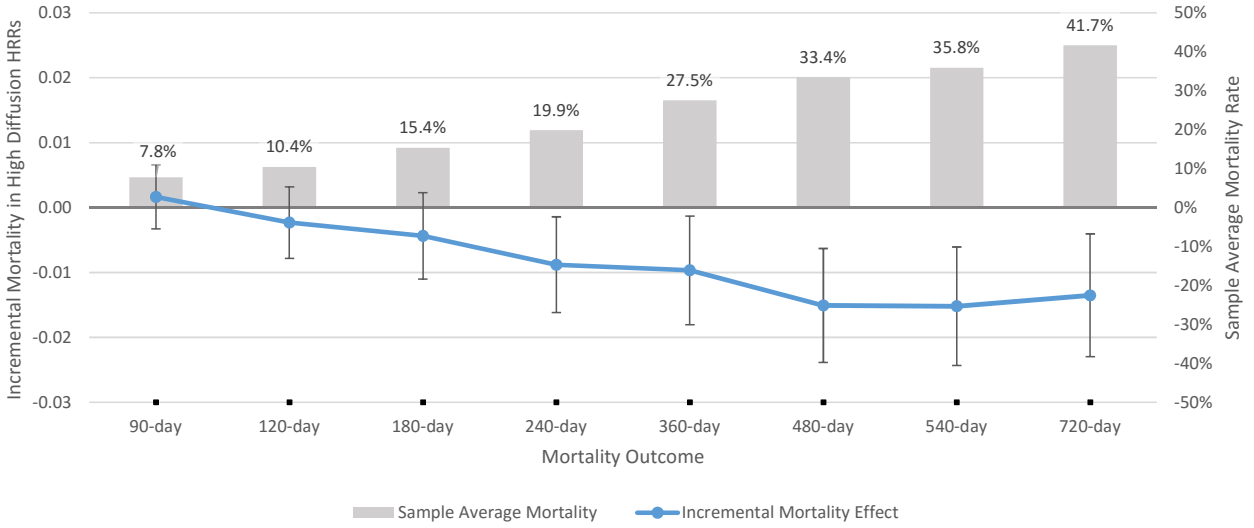
We find no evidence of higher use of the drug for off-label patients in the authors' regions,

[3]For example, capecitabine was initially approved in 1998 for the treatment of metastatic breast cancer that had already proved resistant to both paclitaxel and an anthracycline-containing chemotherapy regimen. In the preceding analysis, we analyzed the use of capecitabine across all breast cancer patients, since our data do not allow us to capture the more specific clinical conditions defining the label indications. However, in the first two years after capecitabine's introduction, a full 39% of its use was on colon cancer patients; colon cancer and breast cancer may be biologically similar, but robust clinical trial evidence was not yet available for the application to colon cancer. The FDA eventually added colon cancer to the label in 2001–after the two-year period covered by our main analysis.

suggesting the authors' influence is largely local to the cancer type on the initial label. A limitation of this analysis is that it is relatively unusual for any given off-label cancer patient to receive treatment with a particular new drug; mean utilization is 0.37 percentage points in regions that ever contain a study author. The point estimate from column (5) suggests that off-label utilization increases by 0.06 percentage points when the first author is in the region; the 95% confidence interval bounds the effect as no larger than 0.16 percentage points.

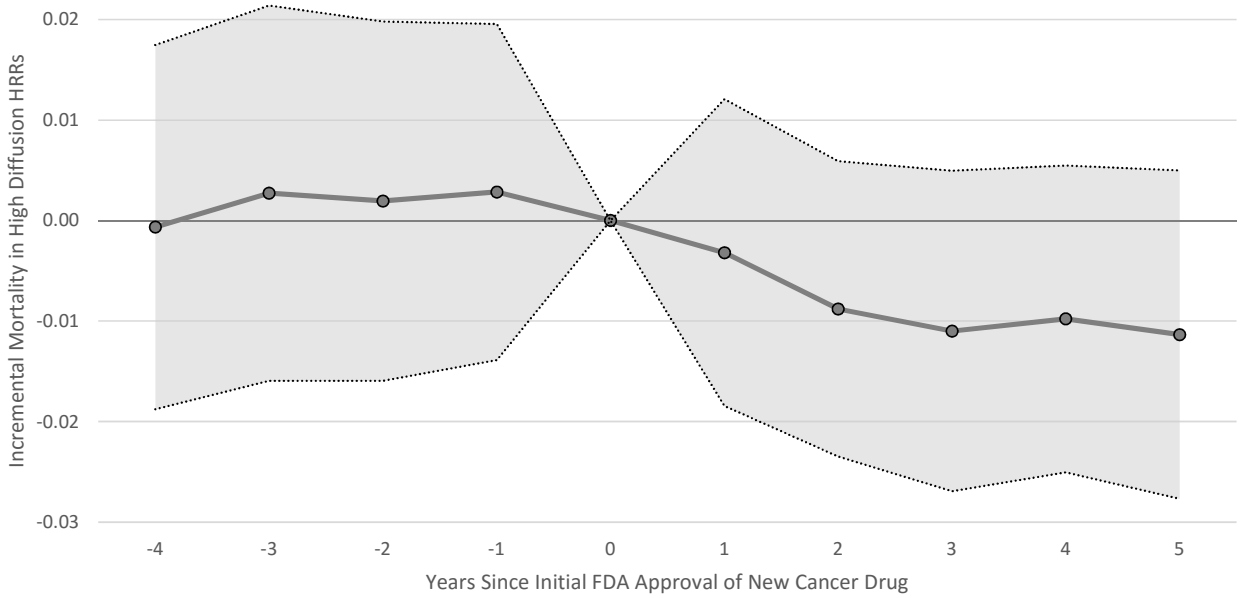
If increased use in the first author's region was driven by a pure "advertising" or "salience" effect boosting awareness or enthusiasm of the new drug, we might have expected greater spillovers to off-label applications. On the other hand, the authors' expertise may be local to the indications studied in the clinical trial they led; they may not have a strong informational advantage when it comes to applications of the drug beyond that population. Taken as a whole, the evidence in this section suggests that first authors boost on-label drug use within their own region, especially when they are located in a region of relative technology laggards, but that authors have little measurable influence on use in neighboring regions or on applications of the drug to other populations.

Figure A1: Mortality Effect of Regional Adoption Speed



Notes: On the primary axis, this solid blue line plots difference-in-differences estimates from eight separate regressions (described in Equation (1)) where the mortality outcome varies from 90-day to 720-day mortality. Each mortality outcome is an indicator for whether a patient died within the specified period from the initial date of the cancer treatment episode. The regression is estimated over new patient-drug episodes that fall within four years before and after FDA approval of the drug. Bands indicate 95% confidence intervals constructed from standard errors clustered at the HRR-drug level. Each difference-in-differences coefficient describes how the mortality rate changes following new cancer drug introductions for indicated cancer patients residing in fast-diffusing regions, compared to mortality rate changes in slow diffusing regions. Regional drug diffusion speed is measured by a leave-out index which is defined as the fraction of cases among all other drugs in our sample for which new drug utilization was higher than the national average over the first four years following FDA approval. This index is a continuous measure ranging from 0 to 1. Thus the plotted coefficients describe the mortality impact of a new drug introduction in places which are above-average early adopters of all other new cancer drugs in our sample (diffusion index = 1) relative to a region which has below-average early utilization of all other sample drugs (index = 0). The secondary axis shows average mortality rates among the regression sample.

Figure A2: Event Study Mortality Effect of Regional Adoption Speed



Notes: This graph plots the coefficients from the difference-in-differences 360-day mortality specification (see Equation (1) and notes to Figure A1), but with the binary "post" indicator replaced by a series of indicators for the number of years relative to the drug's initial FDA approval. The omitted event year is 0. Bands indicate 95% confidence intervals constructed from standard errors clustered at the HRR-drug level. Each coefficient describes how changes in mortality relative to year zero differs across regions with a high versus low propensity to adopt new drugs, where this propensity is measured by a leave-out diffusion index defined as the fraction of cases among all other drugs in our sample for which new drug utilization was higher than the national average over the first four years following FDA approval. Thus estimates of zero in the negative event years prior to drug FDA approval are consistent with the parallel trends assumption that mortality between high- and low-adoption regions would not have changed absent the new drug. Negative mortality estimates in event years following FDA approval suggest larger mortality reductions in fast-diffusion regions compared to slow-diffusion regions.

Notes on Appendix Figures A1 and A2

In this appendix, we discuss the mortality findings summarized in Section 5 of the main paper at greater length. First, it is useful to note a few key features of the identification problem for studying mortality effects. First, our analysis will rely on comparing survival patterns across fast- and slow-diffusion regions. Differences in survival across these regions could be driven either by differences in take-up of the new drug or by differing returns to drug use among infra-marginal treated patients.

Second, to understand the marginal returns to the additional drug use driven by investigator influence, we would ideally compare changes in survival rates across author and non-author regions, echoing our identification strategy from the earlier analysis. The problem is that such comparisons are underpowered for estimating mortality effects: we are relying on moderate variation in drug use from a handful of regions to measure changes in a noisy outcome variable (i.e. mortality rates).

Any effect on mortality would be smaller than the total change in drug utilization, and thus the regression would likely have less statistical power than the main specifications.

Studying mortality rates rather than drug take-up confers one crucial benefit which contributes to identification: we can observe mortality rates for each cancer diagnosis in each region both before and after drug approval in order to construct difference-in-differences estimates of survival effects. Controlling for pre-period differences in mortality allows us to remove regional variation in survival that may be driven by patient health status and isolate survival gains that may be plausibly related to drug introduction. By contrast, our earlier analysis could not exploit such an approach since a new drug's measurable use is zero for all regions in the years before FDA approval and is therefore uninformative about the region's enthusiasm for new technology. For our analysis of drug take-up, we instead relied on regional adoption rates of other new drugs to construct counterfactual adoption behavior.

In light of these considerations, we proceed with a difference-in-differences estimation strategy that compares mortality rates before and after new drug introduction, and across regions with high and low propensities to adopt new cancer drugs. This approach exploits wider variation in regional drug use, while still controlling for any baseline differences in survival by region and cancer diagnosis. Our primary regression specification takes the following form:

$$\begin{aligned}\mathbf{1}(\textit{mortality})_{ijtd} = \beta(\textit{HRR diffusion index})_{jd} \times \mathbf{1}(\textit{post FDA approval})_{jtd} \\ + \{\textit{HRR} \times \textit{drug FEs}\}_{jd} + \{\textit{drug} \times \textit{year FEs}\}_{dt} + \epsilon_{ijtd}. \qquad (1)\end{aligned}$$

The mortality outcome $\mathbf{1}(\textit{mortality})$ is an indicator for whether a patient died within a specified period (e.g. 1-year) of the initial date of the cancer treatment episode. To avoid estimating this relationship using multiple observations from surviving patients, we include only new cancer treatment episodes (i.e. patients with no cancer claims in the previous calendar year). The regression is estimated over patient-drug episodes that fall within four years before and after FDA approval of the drug.

The key coefficient of interest is β which describes how the mortality rate changes after drug introduction in regions that are likely to be fast adopters of the new drug, among patients treated for the relevant cancer type. The variable (*HRR diffusion index*) codes a leave-out estimate of the region's enthusiasm for new cancer drugs, excluding the region's utilization for the particular drug under analysis. To remove potential bias from changes to patient sorting after drug

introduction, we match patients to regions based on their HRR of residence. The indicator variable $\mathbf{1}(\textit{post FDA approval})$ equals 1 in the years following initial FDA approval for drug j and 0 otherwise.

More specifically, $(\textit{HRR diffusion index})_{jd}$ codes for region j the fraction of cases among all drugs in our sample excluding drug d for which new drug utilization in region j was higher than the national average over the first four years following FDA approval. A value of 0 would indicate the region is never an above-average adopter of any other new cancer drug in our sample over the first four years; a value of 1 would indicate the region is an above-average adopter of all other new cancer drugs in our sample. Because we use a leave-out estimate of diffusion speed, this measure should be uncorrelated with the idiosyncratic latent demand for this particular cancer drug in each region, which in turn could be related to differences in health status.

The regression also controls for region by drug (i.e. drug-specific diagnoses) fixed effects. These fixed effects provide a fine level of control for baseline disease-specific regional mortality rates; they are identified in this regression by the inclusion of pre-period years prior to drug introduction. Finally, we control for drug by year fixed effects, which capture national trends in survival for each drug-specific set of cancer diagnoses in our sample.

The identifying assumption for interpreting β as the causal impact of differences in drug utilization on patient survival depends on the usual parallel trends assumption. In this case, the assumption is that fast and slow diffusion HRRs would have experienced parallel changes in mortality rates if no new drug had been introduced. To explore the validity of this assumption, we show the event study plot reporting regression results that replace the binary $\mathbf{1}(\textit{post FDA approval})$ variable with a series of dummy variables for each year relative to FDA approval. An absence of pre-trends in this plot would support the validity of the difference-in-differences results.

Our baseline results from estimating equation (1) are reported in Figure A1. Each point in the figure represents the coefficient β estimated from a separate regression, where the outcome variable is mortality over the indicated time window, ranging from 90 days to 720 days. Beginning with the outcome of 240 day mortality rates, all regressions show statistically significant mortality reductions in fast-adopting regions relative to slow-adopting regions, significant at the 5% level. For example, the point estimate implies a 1 percentage point greater reduction in 1-year mortality for the relevant cancer diagnosis for a region that is an above-average adopter of all other in-sample drugs compared to a region that is a below-average adopter of all other drugs, from a base mortality rate of 27.5%.

To investigate the validity of the parallel trends assumption, we create an event study plot showing the year-by-year differences in 1-year mortality rates across fast- and slow-adopting regions. As Figure A2 shows, the differences in cancer mortality were stable across regions in the pre-period years; after a new drug is introduced, the fast-adopting regions experience steep declines in patient mortality, consistent with the basic difference-in-differences results. These findings corroborate the assumption that fast- and slow-adopting regions were on similar trends in cancer mortality rates before new drug introduction.

To further interpret the magnitude of these results, we examine how our leave-out estimate of regional diffusion speed relates to average new drug utilization rates. To proceed, we use an analogous specification as in (1) above, with pre-period utilization of a new drug before set to zero prior to FDA approval. We find that new drug use is 3.0 percentage points (standard error of 0.36) higher in fast-diffusing regions in the first four years following FDA approval. This change corresponds to a 37% increase in drug utilization over the average rate of 8.1% in the new cancer patient sample. If we interpret the effect on drug utilization as the first stage of an instrumental variables regression to calculate the local average treatment effect of new drug use, the Wald instrumental variable estimate would suggest that new drug use is associated with a 33 percentage point reduction in 1-year mortality, among the marginal treated patients in high-diffusing regions.