Production and Operations Management

Learning from COVID-19: Hospital IT, Clinical Trials and Improvement in Health Outcomes

During the COVID-19 pandemic, US deaths per case decreased from 7.46% in April 2020 to 1.76% in April 2021. In addition to increased testing, a leading explanation for this decline is the effect of hospitals learning that results in better treatment protocols over time for patients diagnosed with COVID-19. Hospitals use health information technologies (IT) to develop digital capabilities that enable greater learning of best practices based on patient health information sharing across providers. In this research, we study whether hospital IT improved patient mortality rates by accelerating the rate at which hospitals were able to learn best practices. Using county-level data on health IT, COVID-19 cases and deaths, we show that counties with greater hospital IT capabilities exhibit fewer COVID-19 deaths. Consistent with the learning effect hypothesis, we show that high IT counties are significantly better at treating COVID-19 cases several months into the pandemic, compared to low IT counties. Counties with hospitals that participated in COVID-19 clinical trials also experienced faster learning. Using LASSO regressions, we find that counties with high IT capability hospitals learned faster through knowledge sharing and greater experimentation, while low IT counties learned only based on their prior experience obtained by treating COVID-19 patients. We posit that hospital IT and clinical trials help hospitals achieve better health outcomes in the long run by enhancing learning effects.

Key words: health IT, hospital IT, COVID-19, learning, digital capability

---

1 Introduction

In March of 2020, public heath experts recommended restricting social contact in order to ‘flatten the curve’ of COVID-19 transmission. In compliance with this guidance, the US federal government announced ‘Fifteen Days to Slow the Spread’ (later, ‘thirty’ days). These guidelines, and the state and local rules that followed them, greatly reduced social mobility, infection rates, and economic output (Chetty et al. 2020). While social distancing has direct economic costs, there is strong reason to think that this strategy was effective in increasing the average level of care received by COVID-19 patients. As observed in Figure 1, US deaths per detected case of COVID decreased from 7.46% in April 2020 to 1.76% in April 2021. If the reduction in death rate can be attributed partly to doctors learning how to better treat COVID-19 patients, then understanding how this process can be accelerated is of utmost importance – both to save lives and reduce the need for economically costly non-pharmaceutical interventions.

In this research, we investigate the role of health information technology (IT) and clinical trials (CT) in the prevention of deaths from COVID-19. We focus on health IT and CTs because they represent natural investments that can help hospitals to develop and deploy better treatments in response to a pandemic of initially unknown etiology. A good electronic health records (EHR) system might may

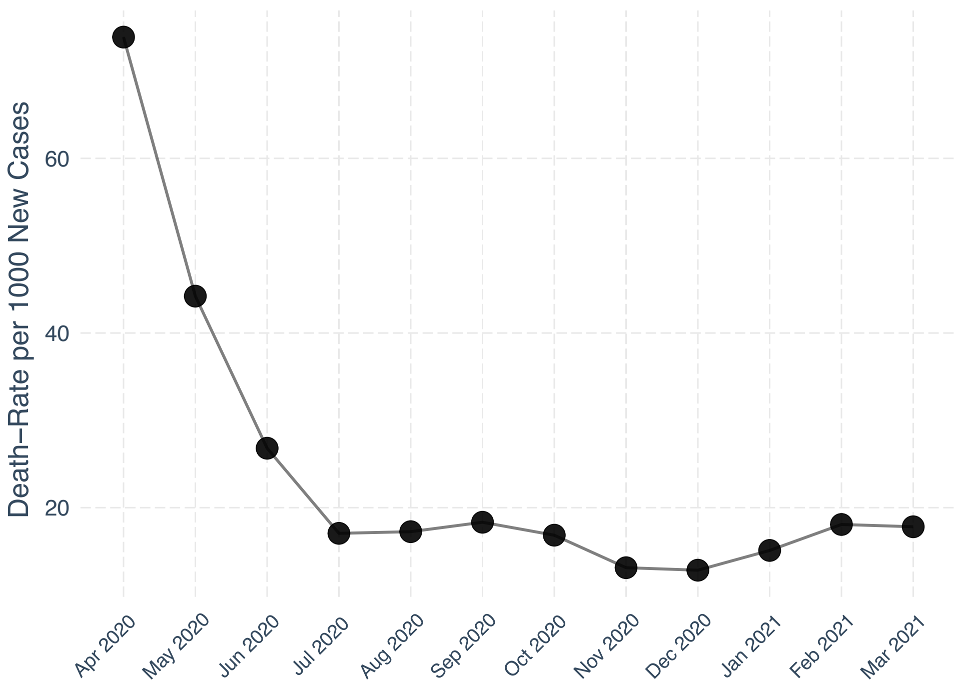

Figure 1 The overall death-rate per thousand COVID-19 cases in the U.S. by calendar month. Death-rate measured as new deaths divided by new cases lagged fifteen days.

help doctors to collect, organize, and share health data on patient diagnosis and outcomes, and is vital for syndromic surveillance during the COVID-19 pandemic. Clinical decision support systems (CDSS) can provide best practice alerts that keep doctors informed on the most cutting-edge modalities for patient care. Similarly, health information exchange (HIE) capabilities that exist in many counties, allow healthcare providers to share patient health data with one another, which enables timely diagnosis and treatment and helps to slow the spread of COVID-19. Further, hospital participation in CTs can also help healthcare providers to learn through experimentation, and develop better understanding of best practices for patient treatment.

We utilize county-level data on Hospital IT intensity, COVID-19 cases, and COVID-19 deaths. Our primary outcome is the COVID mortality rate, measured both as a share of the population and as a share of detected cases. The relationship between COVID-19 deaths per capita and county-level health IT intensity is significant and negative. In our baseline regressions, we find that a standard deviation increase in the county-level index of IT intensity is related to 0.18 fewer deaths per thousand residents. This significant relationship holds after adding fixed effects for state and the number of months into the pandemic. Our preferred specification includes these fixed effects and generates an estimate of .095 (S.E. 0.02) fewer COVID-19 deaths per 1000 residents. As shown in the Appendix A, we attempt to confirm this result as causal by using two approaches to deal with potential endogeneity. The first is to include a large number of additional county-level control variables to address endogeneity concerns attributed to bias due to omitted variables. The second is to instrument for county-level health IT intensity to address concerns related to selection bias and reverse causality. Both approaches provide qualitatively similar results.

We further examine how death rates change as a function of health IT intensity and find that high health IT intensity counties outperform low IT intensity counties in the long-run. However, high IT counties actually do **worse** than others in the first month of the pandemic. Our empirical results suggest that counties with high IT-capability hospitals learned faster in terms of improving patient treatments and best practices over time. Further, we observe that hospital participation in CTs resulted in learning through experimentation, which is associated with lower mortality rates over time. Although IT-intensive counties were poor performers initially, these counties outperformed their low IT-capability counterparts in the long run, thereby saving lives. The poor performance of high IT intensity counties during the initial months may be attributed, in part, to a higher rate of experimentation associated with CTs at hospitals in these counties. This is consistent with changing patient care guidance that ultimately rejected some promising drug candidates (such as Hydroxychloroquine) as counterproductive.

We close by using a LASSO (least absolute shrinkage and selection operator) analysis to select models of learning that predict well out of sample. We then run an OLS regression on the LASSO selected coefficients. The analysis finds that all counties do better over time, but that counties with good clinical trials and those with good IT learn even faster in periods when those specific counties are experiencing the pandemic. Counties without clinical trials fail to learn from additional patients, while firms with low IT do learn from patients, but struggle to hold on to this information over time. In one version of the analysis, firms with clinical trials do worse early in the pandemic, however. This phenomenon of learning through experimentation and knowledge sharing led high IT counties to 'fail faster' initially but achieve stronger long-run results, as manifested by a dramatic decline in COVID-19 mortality rates after the initial months of the pandemic.

2 Background and Hypotheses

In this section, we first provide a background on updates to clinical knowledge about COVID-19 and then develop our research hypotheses on the impact of health IT and clinical trials on patient health outcomes through organizational learning effects.

2.1 Clinical Knowledge of COVID-19

One of the reasons that learning and diffusion of new treatment modalities is important during the COVID-19 pandemic is the rapidly changing state of knowledge on COVID-19. Appendix Figure B.1 provides a timeline of updates to the Infectious Disease Society of America's (ISDA) recommendations for the treatment of COVID-19 patients, with other events related to the spread of the COVID-19 pandemic also included for context. The timeline shows gradual accumulation of knowledge about best practices for treatment for COVID-19. We note that ISDA recommendations are lagging indicators of clinical best practices, as changes in guidelines were often based on trials that were one or two months old during the crisis. On the other hand, the guidelines are leading indicators for doctors that do not closely follow the emerging science and are hesitant to change their treatments. Although doctors are not required to follow these recommendations, they are used to update clinical decision support systems (after a lag) and as a point of reference for malpractice cases. Similarly, even before FDA approval, drugs such as Hydroxychloroquine and Remdesivir may have been prescribed and used off-label. Still, the timeline highlights how perceptions have changed about the classification of drugs in particular.

Hydroxychloroquine, a steroid previously indicated for arthritis and malaria, is the most well known example of a treatment whose public perceptions have changed. On March 19th, 2020 it was hyped as a “game changer” by President Trump, and on March 28th, authorized for emergency use by the FDA. In the first version of the ISDA guidelines, Hydroxychloroquine was presented as a promising but unproven treatment, recommended only in the context of clinical trials. A large scale NIH study duly began on April 2, 2020. On June 20, 2020, the NIH ended its study, stating that the drug was “very unlikely to be beneficial” (Self et al. 2020). Two months later, the ISDA officially recommended against COVID-19 treatment with Hydroxychloroquine (alone or with the antibiotic azithromycin) in version 3.0 of its recommendations.

There have also been changes in perceptions of drugs that were less politically charged. Tocilizumab, another drug previously used for rheumatoid arthritis sufferers, went from a promising drug for further investigation in early ISDA reports to being recommended against in version 3.3 of ISDA’s report. On the positive side, Remdesivir (an antiviral drug previously investigated for use during the 2014 Ebola epidemic) and a few varieties of corticosteroids, such as Dexamethasone, went from promising possibilities to being strongly recommended for specific use cases by the ISDA. There were other developments related to non-drug interventions, such as using prone-positioning to help patients breathe and delaying the use of ventilators, that were also part of these ISDA guidelines.

A recent report on declining COVID-19 death rates confirms a key role for ‘hard-won experience’ at hospitals that have had more time to learn about the disease (Ledford 2020). One doctor emphasized learning better “how to use steroids and a shift away from unproven drugs and procedures.” Other doctors confirmed the important role that uncertainty and over-treatment played in the initial poor response to the disease. “Unfortunately, a lot of initial discourse was complicated by noise about how this disease was entirely different or entirely new” said one quoted doctor, and another agreed that “it took time to realize that standard treatments were among the most effective.” Hydroxychloroquine and Tocilizumab were both mentioned in particular as innovative treatments that were eventually discarded as counterproductive.

Health IT plays an important role in ensuring timely and accurate dissemination of information on best practices across healthcare providers. Specifically, hospitals that develop health information sharing capabilities are more likely to share updated information on clinical best practices and treatment regimens for COVID-19 patients, which is especially important if patients are treated by multiple providers across different facilities (Ayabakan et al. 2017, Janakiraman et al. 2021). Hence, we argue that the ability to share patient health information is a critical measure of hospital IT intensity that is likely to have an impact on patients’ health outcomes.

2.2 Research Hypotheses

As in other industrial sectors, we observe a general trend in implementing IT that supports healthcare operations (Sharma et al. 2016, Bayo-Moriones et al. 2017). For example, EHR systems allow physicians and nurses to manage patient health data efficiently and share patient health information across providers, while telehealth can reduce waiting times for patients and reduce disparities in access to healthcare (Ayabakan et al. 2021).

When used in the context of supporting healthcare operations (such as during surgery or scheduling patients), health IT can produce significant results in terms of patient safety improvements and reduction in operational failures (Tucker et al. 2007). A recent study by Wani and Malhotra (Wani and Malhotra 2018) demonstrates that meaningful use of EHRs improves patient health outcomes. Meaningful use of health IT includes capturing patient information electronically, using patient information to track key clinical conditions, integrating test and imaging results, using clinical decision support tools, and communicating patient health information to all providers involved in patient care for care coordination (Wani and Malhotra 2018). Health IT supports better decision making by facilitating integration of patient medical history for safe and relevant diagnosis in a timely manner (Kohli and Tan 2016). Hence, health IT systems can support greater learning through two mechanisms: (a) diffusion of new knowledge on best practices and treatments for COVID-19 patients between clinicians, and (b) enabling greater experimentation through clinical trials that foster a culture of learning through trial and error.

Clinicians can use health IT to develop greater organizational capabilities in case management, patient engagement, post-acute follow up, health information sharing, telehealth and remote patient monitoring. Recent studies have explored the use of health IT systems in improving patient care coordination and transitions across provider settings, by facilitating effective patient assessment and discharge planning (Ayabakan et al. 2021). CDSS's improve decision-making capabilities by identifying drug-drug and drug-allergy contraindications, thereby reducing the incidence of adverse drug events and yielding significant savings in inpatient care (Amarasingham et al. 2009). CDSS's also aid in short-term preventive care and chronic disease management and syndromic surveillance of infectious diseases. For example, heuristics within EHR systems can identify patients who are at high-risk, remind physicians to order needed tests and schedule preventive care visits, thereby leading to better clinical outcomes such as lower mortality and readmission risk.

In the context of a pandemic, patient mobility creates a fragmented environment where healthcare services are delivered by multiple, often unaffiliated providers. This poses a challenge for communication among caregivers necessary for care coordination across the provider network. The availability of timely and accurate health records about patients' medical histories, diagnosis, test results, and medications, is key to controlling the spread of the COVID-19 virus and reducing the number of deaths. Lack of information sharing about patient diagnosis and health data can lead to medical errors (Atasoy et al. 2018). Such adverse events may extend patients' length of hospital stay and also increase the risk of hospital-acquired infections and mortality. Sharing patient health data also allows providers to continuously monitor disease progression, which facilitates better management of available hospital resources such as personal protective equipment, ICU beds, and other constrained resources such as ventilators. Hence, hospital participation in health information sharing initiatives, such as state-level and nation-wide information exchanges as well as disease and immunization registries, are likely to improve digital capabilities to counter the spread of a pandemic. At a conference of the American Medical Informatics Association on Nov 17, 2020, Dr. Anthony Fauci stressed the importance of data sharing and tracking as integral parts of the COVID-19 response, especially with respect to tracking and isolating COVID-19 patients, contact tracing and isolation/quarantining[1]. "*There are so many things that we need better data on,*" said Fauci. "*It's there. The question is collecting it, putting it in a form that can be distributed.*"

Based on the critical role of health IT in enhancing learning effects through knowledge sharing and experimentation, we posit that greater use of health IT can provide hospitals with digital capabilities necessary to improve learning and thereby reduce mortality rates due to COVID-19. Hence, we hypothesize that,

H1: *Counties with high health IT capabilities exhibit lower mortality rates due to COVID-19, compared to low health IT counties.*

2.2.1 Learning Effects

The benefits of learning from experience have received significant attention in the operations literature based on the widely held belief that "practice makes perfect." A large body of empirical research has documented the association between cumulative experience and performance improvement, providing an empirical basis for the concept of learning-by-doing (Von Hippel and Tyre 1995). However, performance improvements may be attributed to the cumulative experience gained (i.e., an experience effect) without considering the role of organizational learning (i.e., an organizational learning effect) (Roth and Menor 2003).

When organizations perform certain tasks repeatedly, they learn, develop and adapt routines to their needs and environment (Nelson and Winter 1982). Organizational learning is a process of seeking, selecting, amending, and adapting new routines to improve performance. Organizational learning accumulates over time through repetitive execution (Pisano et al. 2001). Participants contribute to the learning process by observing, doing, reforming, and sharing lessons learned with each other. Without organizational learning, cumulative experience alone cannot guarantee significant performance improvement (Tucker et al. 2007).

Compared to hospitals that do not participate in clinical trials, hospitals with a higher rate of CT participation are more likely to develop relevant process and organizational routines to care for COVID-19 patients, by developing critical care pathways and fostering greater involvement with case managers (Sosa et al. 1998). Further, high CT hospitals are more likely to comply with established routines for care management of COVID-19 patients, ensuring effective health care delivery based on their case

[1] See `https://tinyurl.com/y4la93re`, accessed on Nov 30, 2020.

experience. Zollo and Winter (Zollo and Winter 2002) argue that high-frequency practices facilitate additional capability building. Hence, hospitals with greater participation in clinical trials are more likely to have the relevant experience in dealing with challenges involved in treating COVID-19 patients across transitions in care settings, due to the cumulative learning effects in dealing with patients with a diverse array of comorbidities and treatments. Hence, we hypothesize that,

H2: *Counties with hospitals that participated in clinical trials exhibit lower patient mortality rates due to learning effects, compared to counties that did not participate in clinical trials.*

While learning is generally viewed as greatly dependent on cumulative organizational experience, some studies have indicated that the effect of past experience may depreciate over time (Benkard 2000), thereby highlighting the importance of studying the role of recency of experience on patient health outcomes (Huckman and Pisano 2006). The basic mechanism through which prior experience affects individual performance is learning through knowledge transfer (Avgerinos and Gokpinar 2018). Further, surgeons typically make several decisions during surgery and tend to respond based on past stimuli and learning from prior experience (Loewenstein and Lerner 2003). Therefore, priming from prior experience can help surgeons make better decisions during surgery to improve patient outcomes. Although COVID-19 patients may have different medical needs based on their prior medical history and vulnerabilities, COVID-19 treatment protocols share many similarities based on clinical guidelines. Hospitals with greater health IT capabilities are more likely to share COVID-19 treatment protocols based on the experience of clinicians and healthcare providers in treating COVID-19 patients (Ayabakan et al. 2017). In other words, health IT can facilitate sharing of best practices on COVID-19 treatments across providers within and across different health organizations, and thereby allow healthcare providers to benefit from the collective knowledge and experience of others through a knowledge spillover effect.

On the other hand, hospitals with low IT capabilities lack the infrastructure necessary to share patient health information and treatment protocols across different providers, resulting in higher costs and poor health outcomes (Atasoy et al. 2018, Janakiraman et al. 2021). Such hospitals are dependent on learning through in-house experience based on experimentation through clinical trials. In other words, low IT intensity hospitals learn through trial and error based on their own recent experience in dealing with COVID-19 patients and clinical trials that are conducted within their own organizations (i.e., an experience effect). However, a lack of adequate IT capabilities precludes them from learning through knowledge spillover from other providers and healthcare delivery organizations. Hence, we argue that learning effects due to recent experience (in treating COVID-19 patients) is relatively more important in counties with low IT-intensity hospitals, while high IT-intensity counties are more likely to benefit from learning through knowledge sharing and spillover from other healthcare providers.

H3: *Recent experience with COVID-19 cases has a greater learning effect on mortality rates in low health IT counties, while high health IT counties exhibit greater learning effects due to knowledge sharing (i.e., a spillover effect).*

3 Data

Our study data falls into four categories. Our primary outcome of interest is the number of deaths attributed to COVID-19 which we measure monthly at the county-level (i.e., the county-level mortality rate). We measure COVID cases at the same level of resolution. We study whether and how hospital IT intensity makes their counties resilient to COVID-19, and hence, the individual measures of hospital IT intensity comprise our second data set. We also gather data on clinical trials being conducted at hospitals in the county. Finally, we collected a battery of covariates and instruments for our robustness analysis. We will discuss these variables and the data collection process in the next section.

3.1 COVID Death and Case Data

Our COVID cases and deaths data was obtained from the data repository curated by the Center for Systems Science and Engineering at Johns Hopkins University (Dong et al. 2020). They obtain their data from numerous resources including the US Centers for Disease Control and local public health agencies. The data is available daily at the county level. Figure 2 plots the distribution of cumulative cases and deaths across all counties in the US as of April 5, 2021. These plots reflect the long-tail nature of the distribution which shows that approximately 50 counties accounted for a disproportionate share of the COVID-19 cases and deaths in the United States.

Our primary dependent variable is deaths per capita, which we define as the cumulative number of deaths in a county, from the advent of COVID-19 until the end of our study time period (i.e., April 5, 2021) divided by the population of that county. Additionally, we also considered the cumulative number of deaths as a ratio of the cumulative cases, as an alternative dependent variable and obtain qualitatively similar results. The other dependent variable that we utilize to test our learning hypothesis (H2) is the ratio of the number of new deaths in a county to the number of new cases, lagged by two weeks. Cases are lagged by two weeks since a fatality attributed to COVID-19 is typically detected as a positive case approximately two weeks before death (Kraemer et al. 2020, Flaxman et al. 2020).[2] Our primary measure of time is number of months elapsed since the first 10 COVID-19 cases detected in a county. Counties with a longer elapsed time since initial exposure will have higher cumulative deaths because they were exposed to the pandemic for a longer period.

3.2 Defining County Level Hospital IT Measures

Our data on hospital IT is sourced from the Healthcare Information and Management Systems Society (HIMSS) Electronic Medical Record Adoption Model (EMRAM) in 2020.[3] HIMSS data is widely used in the health IT literature to measure IT adoption by hospitals (for e.g. Kwon and Johnson (2014), Lee et al. (2013)).

[2] Our results are robust to dropping this lag and to lags of up to 3 weeks.

[3] https://www.himss.org/what-we-do-solutions/digital-health-transformation/maturity-models/electronic-medical-record-adoption-model-emram; accessed on Oct 31, 2021.

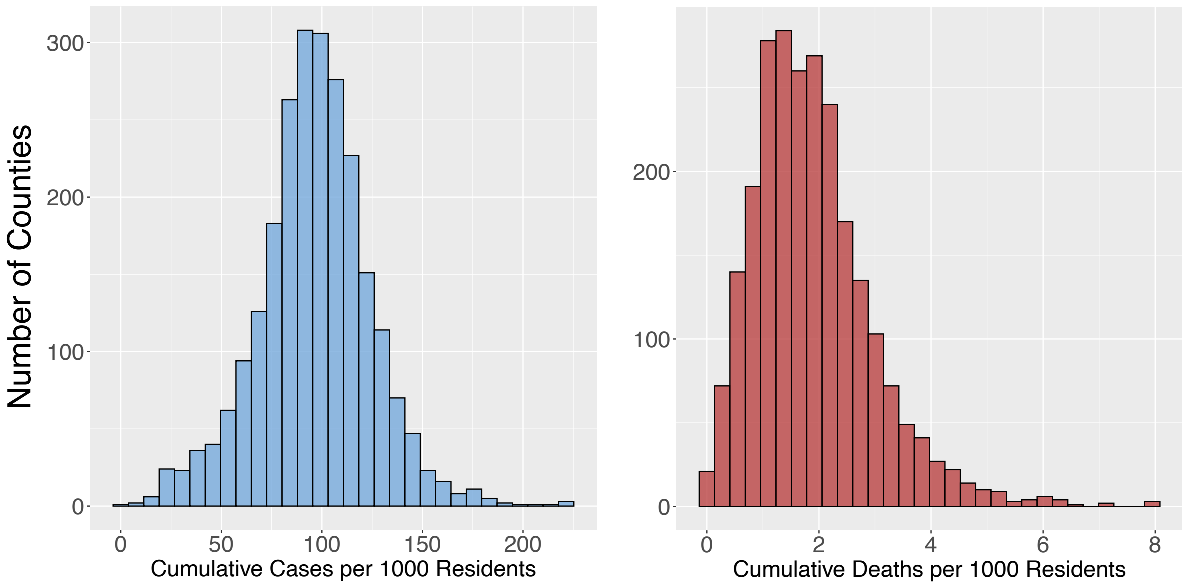

Figure 2 (A) Histogram of cumulative cases per 1000 individuals in a county as recorded on Apr 5, 2021. (B) Histogram of cumulative deaths per 1000 individuals in a county as recorded on Apr 5, 2021.

The EMRAM is a widely accepted industry standard that is used to measure the maturity of a hospital's EHR implementation and adoption of health IT to support patient care, reduce medication errors, and optimize operational throughput, in a paperless environment. Every hospital is assigned to a "stage" ranging from 0 - 7 depending on their current level of overall hospital IT implementation and adoption. If a hospital is assigned a rating in stage 6 or 7, we consider them to have high levels of hospital IT maturity. Hospitals in the "Advanced EMRAM" category (in stage 6 or 7) are likely to have implemented and actively use advanced clinical decision support systems, technology-enabled medication reconciliation, risk reporting, complete electronic medical record adoption (no paper), data analytics, and participate in electronic health information exchanges (where patient data is shared with other hospitals). Hence, the EMRAM categories represent different dimensions of hospital digital capability and influence how hospitals may react to health emergencies including spread of the COVID-19 pandemic.

Our key treatment variable is a health IT index which measures the IT intensity of all hospitals within a county. To construct this index, we first assign a binary variable if a hospital is evaluated to be at stage 6 or 7 of the EMRAM model. To convert our hospital IT metric into a county-level health IT index, we used the normalized weighted average of hospital IT indices across all hospitals in a county weighted by their number of admissions (based on HIMSS annual hospital survey data). Our key independent variable, "Health IT index," represents a county-level measure of health IT capability across all hospitals in the county, weighted by hospital size. Although this measure does not include IT use across other healthcare providers (such as nursing homes, physician offices, etc.), we believe that it represents a useful starting point to study the impact of health IT, especially since most severe cases of COVID-19 are treated at a hospital during at some time during their care.

3.3 Clinical Trial Data

Data on clinical trials related to COVID-19 were retrieved from clinicaltrials.gov on March 8, 2021. There were 4968 clinical trials in the database related to COVID-19 first posted on or before this date. Information on the database is provided by the sponsor or PI of the clinical study. Federal and state laws require many types of clinical trials to be registered with this database, but the data is not strictly comprehensive (ClinicalTrials.gov 2021).

Clinical trials were classified into interventional or non-interventional trials, based on their descriptions. Interventional clinical trials were further classified into treatments judged to be Good, Bad, or Mixed/Neutral based on the best available judgment on their efficacy as of April 2021. When more than one type of clinical trial was present in the county, we code it as using the best of the ongoing varieties.[4]

3.4 Complementary Data Sources

We utilize the 2018 American Community Survey (ACS) data to include a rich set of demographic (age, gender, race) and economic variables (income, income inequality, employment composition, employment density) as controls in our econometric models. ACS is a nationally representative survey conducted by the US census and sent to over 3.5 million households across the US every year. We complement this data with Rural-Urban Continuum Codes (RUCC) from the Economic Research Service of the US Department of Agriculture.[5] These codes distinguish counties into nine categories based on their degree of urbanization and connection to urban centers.

To control for pre-Covid hospital quality, we use data on hospital level mortality from the Centers for Medicare Medicaid Services (CMS). To control for mobility in a county, we use data from Google's COVID-19 community mobility reports.[6] Using location data from Android phones, Google calculates the relative change in population mobility across various types of locations in a given county with respect to baseline mobility in January, 2020 (pre-Covid). Using this data, we calculate the average change in county-level mobility relative to the pre-Covid levels of mobility in the same country during the duration of our study period. To control for user mobility inflows into a county (from other counties), we used data from the University of Maryland Transportation Institute's COVID-19 impact analysis platform and calculated total inflows into a county during our study period (Zhang et al. 2020). This dataset was aggregated from several different mobile location data providers and previously used to study the impact of mobility on COVID-19 infections (Xiong et al. 2020). We augment these datasets with data on availability of airports in a county (from OurAirports.com) and data on trust in science based information sources at the state level (from the MIT COVID-19 beliefs, behaviors and norms

[4] Our two most important references for best practices in COVID-19 treatment were the ISDA's recommendations (American Journal of Managed Care, Staff of 2020) and the continually updated NYT COVID-19 Drug and Treatment feature `https://www.nytimes.com/interactive/2020/science/coronavirus-drugs-treatments.html`

[5] See `https://www.ers.usda.oducts/rural-urban-continuum-codes/documentation/`, accessed on Nov 22, 2020, for details about RUCC.

[6] Google COVID-19 Community Mobility Reports, `https://www.google.com/covid19/mobility/`, Accessed: April 5, 2021.

survey (Collis et al. 2020)). Tables 1 and 2 provide summary statistics of our dependent and independent variables.

Table 1 Summary statistics of county level time-invariant variables.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Population	2,449	129,874.400	373,764.600	1,232	15,720	97,241	10,039,107
Population Density	2,425	266.849	1,028.369	0.244	22.010	159.823	25,590.650
No of Hospitals	2,449	2.488	4.404	1	1	2	101
Cum. Deaths (Apr 5 2021)	2,449	216.236	764.477	0	26	150	23,301
Cum. Cases (Apr 5 2021)	2,449	12,007.500	39,585.990	0	1,415	8,576	1,222,479
N. of Days since 10 cases	2,449	338.281	54.219	0	319	373	400
Mobility to work (change)	2,246	−23.590	6.200	−67.000	−26.513	−19.454	−7.799
N. of Beds weigh. Emram Index	2,449	−0.000	1.000	−0.749	−0.749	0.991	1.674
% of Males	2,425	0.499	0.020	0.443	0.488	0.504	0.663
% of White Pop.	2,425	0.834	0.155	0.117	0.773	0.947	0.999
% of Black Pop.	2,425	0.089	0.140	0.000	0.008	0.101	0.874
% of Indian Pop.	2,425	0.016	0.054	0.000	0.002	0.008	0.780
% of Asian Pop.	2,425	0.015	0.025	0.000	0.004	0.015	0.359
% of 0-24y Pop.	2,425	0.316	0.045	0.105	0.289	0.337	0.612
% of 25-40 Pop.	2,425	0.179	0.027	0.072	0.161	0.193	0.346
% of 40-65 Pop.	2,425	0.325	0.028	0.159	0.312	0.344	0.408
% of >65 Pop.	2,425	0.180	0.044	0.064	0.152	0.204	0.556
Trust in Science Index	2,449	0.669	0.062	0.357	0.627	0.708	0.875
Median Income	2,425	52.180	13.542	20.188	43.373	57.765	136.268
% of hlds with Broadband	2,353	62.075	10.873	22.500	55.300	69.300	91.300
% workers in IT	2,353	1.452	0.791	0.000	1.000	1.800	15.800
Download Internet speeds	2,446	84.373	41.466	4.905	51.891	115.381	198.289
Employment Density	2,425	144.835	1,468.927	0.000	7.000	57.000	67,846.000
No of Large Airports	2,449	0.065	0.268	0	0	0	3
No of Medium Airports	2,449	0.218	0.499	0	0	0	6
No of Small Airports	2,449	4.528	4.619	0	2	6	51
Inflow from Neigh.Counties	2,093	1,081.821	1,811.403	17.513	230.247	1,098.191	26,808.630

Table 2 Summary statistics of county level time-invariant variables.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
New Cases 1 Month since 10 Cases	3,063	70.111	238.536	1.000	13.000	52.500	6,713.000
New Cases 2 Months since 10 Cases	3,057	311.262	1,814.451	0.000	13.000	107.000	43,087.000
New Cases 3 Months since 10 Cases	3,053	258.050	1,233.000	0.000	20.000	128.000	44,606.000
New Cases 4 Months since 10 Cases	3,055	346.850	1,353.614	0.000	39.000	239.000	38,802.000
New Cases 5 Months since 10 Cases	3,042	633.801	2,940.026	0.000	61.000	371.000	75,753.000
New Cases 6 Months since 10 Cases	3,005	525.722	2,127.485	0.000	72.000	389.000	71,547.000
New Cases 7 Month since 10 Cases	2,951	499.124	1,318.439	0.000	94.000	492.500	39,510.000
New Cases 8 Months since 10 Cases	2,876	765.746	1,562.652	0.000	133.000	809.750	31,134.000
New Cases 9 Months since 10 Cases	2,642	1,581.465	3,851.698	0.000	192.000	1,429.750	101,827.000
New Cases 10 Months since 10 Cases	2,311	2,382.514	7,385.905	0.000	222.000	1,820.500	196,059.000
New Cases 11 Months since 10 Cases	1,994	2,760.921	12,194.390	0.000	183.250	1,816.750	426,426.000
New Cases 12 Months since 10 Cases	1,361	1,910.960	7,031.753	0.000	183.000	1,408.000	202,770.000
New Deaths 1 Month since 10 Cases	3,064	8.973	74.019	0.000	0.000	3.000	2,336.000
New Deaths 2 Months since 10 Cases	3,064	18.429	119.398	0.000	0.000	3.000	2,881.000
New Deaths 3 Months since 10 Cases	3,062	11.172	76.095	0.000	0.000	3.000	2,112.000
New Deaths 4 Months since 10 Cases	3,060	8.221	36.332	0.000	0.000	5.000	1,009.000
New Deaths 5 Months since 10 Cases	3,046	10.412	47.908	0.000	0.000	7.000	1,259.000
New Deaths 6 Months since 10 Cases	3,011	9.306	39.197	0.000	1.000	7.000	1,115.000
New Deaths 7 Month since 10 Cases	2,952	8.984	26.432	0.000	1.000	9.000	792.000
New Deaths 8 Months since 10 Cases	2,881	11.623	21.624	0.000	2.000	14.000	480.000
New Deaths 9 Months since 10 Cases	2,648	21.021	44.195	0.000	2.000	22.000	1,056.000
New Deaths 10 Months since 10 Cases	2,313	32.543	85.263	0.000	3.000	30.000	2,167.000
New Deaths 11 Months since 10 Cases	1,994	39.594	175.245	0.000	3.000	30.000	6,321.000
New Deaths 12 Months since 10 Cases	1,361	35.293	167.838	0.000	3.000	24.000	5,201.000
Death-Rate over 1000 Cases 1 Month since 10 Cases	3,064	49.825	75.157	0.000	0.000	76.923	670.330
Death-Rate over 1000 Cases 2 Months since 10 Cases	3,029	34.354	57.443	0.000	0.000	52.469	526.316
Death-Rate over 1000 Cases 3 Months since 10 Cases	3,034	26.809	46.545	0.000	0.000	38.462	818.182
Death-Rate over 1000 Cases 4 Months since 10 Cases	3,047	22.929	40.645	0.000	0.000	28.571	602.606
Death-Rate over 1000 Cases 5 Months since 10 Cases	3,037	18.917	27.271	0.000	0.000	25.641	500.000
Death-Rate over 1000 Cases 6 Months since 10 Cases	2,999	22.029	30.700	0.000	4.029	28.103	500.000
Death-Rate over 1000 Cases 7 Month since 10 Cases	2,945	22.419	30.579	0.000	5.988	28.455	500.000
Death-Rate over 1000 Cases 8 Months since 10 Cases	2,877	19.563	25.549	0.000	6.780	25.194	500.000
Death-Rate over 1000 Cases 9 Months since 10 Cases	2,640	19.386	25.010	0.000	7.774	23.256	500.000
Death-Rate over 1000 Cases 10 Months since 10 Cases	2,308	20.225	21.806	0.000	9.126	24.727	363.636
Death-Rate over 1000 Cases 11 Months since 10 Cases	1,992	21.368	24.786	0.000	9.819	25.680	500.000
Death-Rate over 1000 Cases 12 Months since 10 Cases	1,359	23.101	23.958	0.000	9.292	28.026	277.778

4 Health IT and COVID-19 Mortality

Did hospital information technology reduce the mortality rate due to COVID-19? To address this central research question, we begin by visually inspecting the data. Figure 3 presents the cumulative number of COVID-19 deaths per capita for US counties, where the observations are binned based on the number of months since the first ten cases were discovered in the county. Counties are split into two categories - low and high IT capability - based on the IT intensity of the hospitals in the county. The figure shows the log-population weighted average of the cumulative deaths over 1000 individuals (along with 95% confidence intervals) for the 12 months since the first 10 recorded cases for the two groups of counties. Although the cumulative number of deaths increases over time across all types of counties, what is most striking is the divergence over time between deaths per capita in counties of different IT capabilities. Specifically, we observe that, during the first month of exposure to COVID-19, counties with different levels of IT capability had similar number of deaths per-capita, even though different counties were exposed to COVID-19 at different times. High IT counties had 0.06 deaths per thousand residents (with [0.056,0.064] 95% confidence interval), while low IT counties had 0.088 deaths per thousand residents (with [0.081,0.096] 95% confidence interval). However, the rate of increase in deaths for counties with high IT hospitals stayed constant or decreased, while counties with low IT hospitals witnessed significantly larger growth in the average rate of deaths per capita. Six months after the first ten cases were detected in a county, high IT counties exhibited COVID-19 deaths totalling 0.518 per thousand residents (with [0.497, 0.539] 95% confidence interval) whereas low IT intensity counties had 0.781 deaths per thousand residents (with [0.747, 0.815] 95% confidence interval).

4.1 Regression Analysis

Next, in order to better understand the relationship between county-level, health IT capability and COVID-19 outcomes, we conduct econometric analyses using ordinary least squares (OLS) regressions. Table 3 reports the results of five variations on the cross-county regression model shown in equation (1).

$$\text{DeathsPerCapita}_c = \beta \text{Health IT-Index}_c + \lambda X_c + \varepsilon_c \tag{1}$$

where c indexes the particular county, and X_c is a vector of county-level controls and fixed effects.

Column (1) of Table 3, which contains no other fixed effects or controls, finds that a standard deviation increase in county-level health IT capability is significantly associated with .180 fewer deaths per thousand residents of population. Adding a control for the number of days since the first ten cases, either with a linear term or with fixed effects for the number of months of exposure, as shown in columns (2) and (3), respectively, only somewhat decreases the size and magnitude of the effect. Adding a state-specific fixed effect in column (4) further attenuates the estimate, although it remains negative and highly significant. In column (5), we further control for other county arrtibutes include Median Income (in thousands), % of African Americans, change in mobility to workplace (Google data), number of

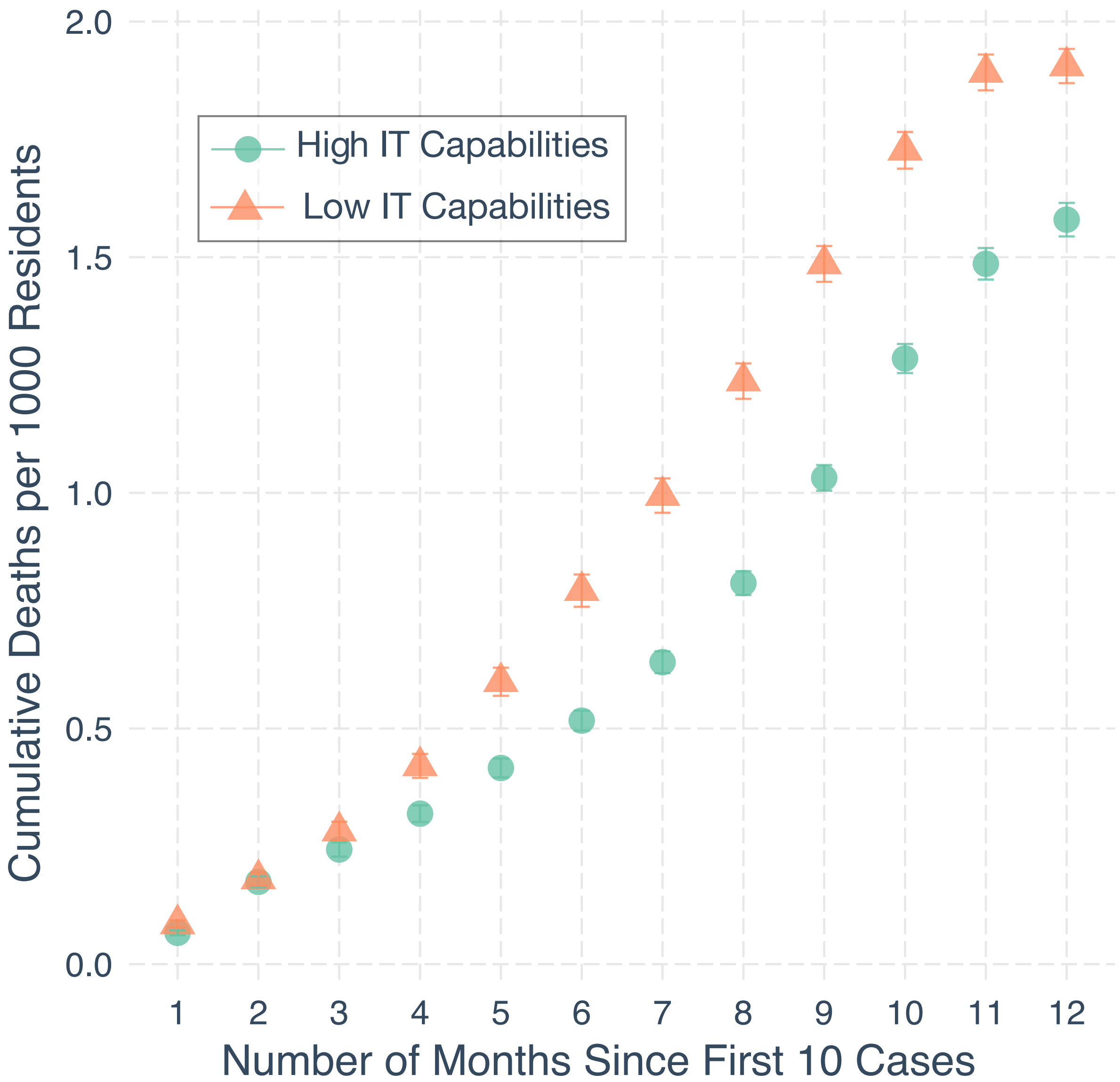

Figure 3 Cumulative COVID-19 deaths per capita as a function of hospital IT index (based on emram index) and months of exposure to the COVID pandemic across US counties.

hospitals (FE) and RUCC (FE). This is our preferred estimate for reporting purposes. That is, after controlling for the first time that a county was exposed to COVID-19 and inclusion of a state fixed effect, and a battery of controls, we find that a standard deviation increase in county-level health IT index is associated with .063 fewer deaths per thousand residents.

The dependent variable in the final column of Table 3 represents an alternate DV. It measures the cumulative number of deaths as a ratio of the cumulative number of cases detected in the county, rather than number of residents. We focus on deaths per capita as our main outcome of interest in columns (1) to (5), because including the number of cases induces an additional level of measurement error relative to deaths.[7] However, switching to death rates does not change our qualitative findings, as observed by

[7] Early in the pandemic, only severe cases of COVID-19 were detected. Wide scale testing means that less fatal cases were detected over time, contributing to a decrease in observed deaths per case.

the significant coefficient of -0.993 of "health IT Index."[8] In other words, after controlling for state-level fixed effects and number of months of exposure, our results indicate that a standard deviation increase in county health IT capability is associated with 0.99 fewer deaths per thousand cases. Our results support H1 with respect to the negative association between high IT capability counties and lower mortality rates.

	Cum Deaths per 1000 Residents					Cum Deaths per 1000 Cum.Cases
	(1)	(2)	(3)	(4)	(5)	(6)
Health IT Index	-0.180^{***} (0.021)	-0.164^{***} (0.036)	-0.156^{***} (0.034)	-0.095^{***} (0.020)	-0.063^{***} (0.018)	-0.993^{**} (0.382)
100 Days since 10 Cases		-0.129 (0.151)				
No of Months since 10 cases (FE)	No	No	Yes	Yes	Yes	Yes
State (FE)	No	No	No	Yes	Yes	Yes
Other Controls*	No	No	No	No	Yes	No
Constant	1.890^{***} (0.021)	2.095^{***} (0.279)				
Observations	2,430	2,430	2,430	2,430	2,215	2,430
R^2	0.028	0.031	0.048	0.361	0.437	0.281
Adjusted R^2	0.028	0.030	0.045	0.345	0.411	0.263
Residual Std. Error	1.051	1.050	1.042	0.863	0.758	18.430
df	2428	2427	2421	2370	2116	2370
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$					

Table 3 OLS Regression Results following equation (1).

While our results indicate a strong negative association between the county health IT index and deaths per capita, this relationship need not be causal. The negative relationship could, in principle, be driven by an omitted variable that causes a county to adopt more health IT and also reduces its vulnerability to deaths attributed to COVID-19. In general, either selection bias or reverse causality could also result in a spurious relationship, but neither factor is plausibly in play here. Our data is based on the complete set of US counties, ruling out selection bias, and it is highly implausible that hospitals made IT investment decisions in the past due to their beliefs about their county's susceptibility to this unprecedented (and unforeseen) pandemic. Such health IT investments typically exhibit a time lag of 12-18 months before they can be implemented and adopted (Dranove et al. 2014). Further, to the extent that reverse causality is a concern, it would tend to bias **against** our findings – hospital administrators would likely want to make **larger** investments in health IT if they anticipate a pandemic.

To further strengthen our claims, we perform several additional robustness checks in Appendix A. These include controlling for several other demographic and county characteristics in Table A.1, controlling for pre-Covid hospital quality in Table A.2 and controlling for work from home and social

[8] In our hypotheses section, we do focus on death rates, because counties experienced waves of COVID-19 cases at different points in time, thereby obscuring the relationship between deaths and knowledge accumulation.

distancing behavior in Table A.3. We also use an instrumental variable analysis where we instrument for health IT quality with measures of broadband internet availability and download speeds in Appendix Table A.4 (first stage) and Table A.5 (second stage). All of these robustness checks produce qualitatively similar results.

5 Health IT and Clinical Learning

We hypothesize that an important pathway through which health IT may help hospitals to prevent COVID-19 fatalities is by helping hospitals to learn, share and implement clinical best practices for treatment of COVID-19 patients. We test this hypothesis (H2) in two ways. We show that counties with higher levels of IT capability in their hospitals improved their treatment and management of COVID-19 cases faster than low-IT capability counties. We then investigate the cause of faster COVID-19 improvement in counties with higher levels of IT-intensive hospitals. The accelerated improvement among high health IT counties is driven by younger, and not older, counties. This rules out the 'dry tinder' explanation, i.e., that early poor-performing counties do better during the pandemic's later stages because much of the vulnerable population may already be dead.

We also find evidence that faster learning among high-IT counties is driven in part by clinical trials. We show that counties with IT-intensive hospitals conducted COVID-19 related clinical trials at a much higher frequency than less IT-intensive counties. In fact, counties in the bottom half of the health IT index accounted for only 0.65% of total clinical trials. Comparing counties with and without COVID-19 trials, we observe the same learning pattern over time: hospitals with more COVID-19 trials had more COVID-19 deaths per case at the onset of the pandemic, but in the long run, they performed better than counties without COVID-19 clinical trials. This is consistent with counties with high IT hospitals aggressively experimenting with innovative treatments that later turned out to be counterproductive.

However, the knowledge gained by hospital staff through this experimentation turned out to be invaluable, both generally and for the participating hospital. As the initial wave of innovative treatments were rejected (due to evidence from clinical trials), and new, better procedures and medications emerged, all hospitals saw their death rates decline. The decline in death rates was particularly dramatic among hospitals with clinical trials, suggesting that some of the knowledge gains from testing were concentrated in the innovating hospital/county, and perhaps that the most recent round of clinical trials (for drugs such as the monoclonal antibody, Regeneron) have had more salutary effects.

5.1 Learning in Action

In a hospital setting, learning about how to better treat COVID-19 patients should be manifested in lower mortality rates over time. To study whether counties with more IT-intensive hospitals learned faster, we need to compare the rate at which their death rates decreased, and not just the total amount of deaths.

Figure 4 presents model-free evidence on this question. It shows the number of deaths per thousand cases (lagged by two weeks which represents the 14-day incubation period for COVID-19 cases), over time, experienced in counties of different types. Counties are categorized into two bins based on their health IT-capability.[9] County averages are weighted by the logarithm of the county population. Time is measured as number of months since the first ten cases were diagnosed in the county.

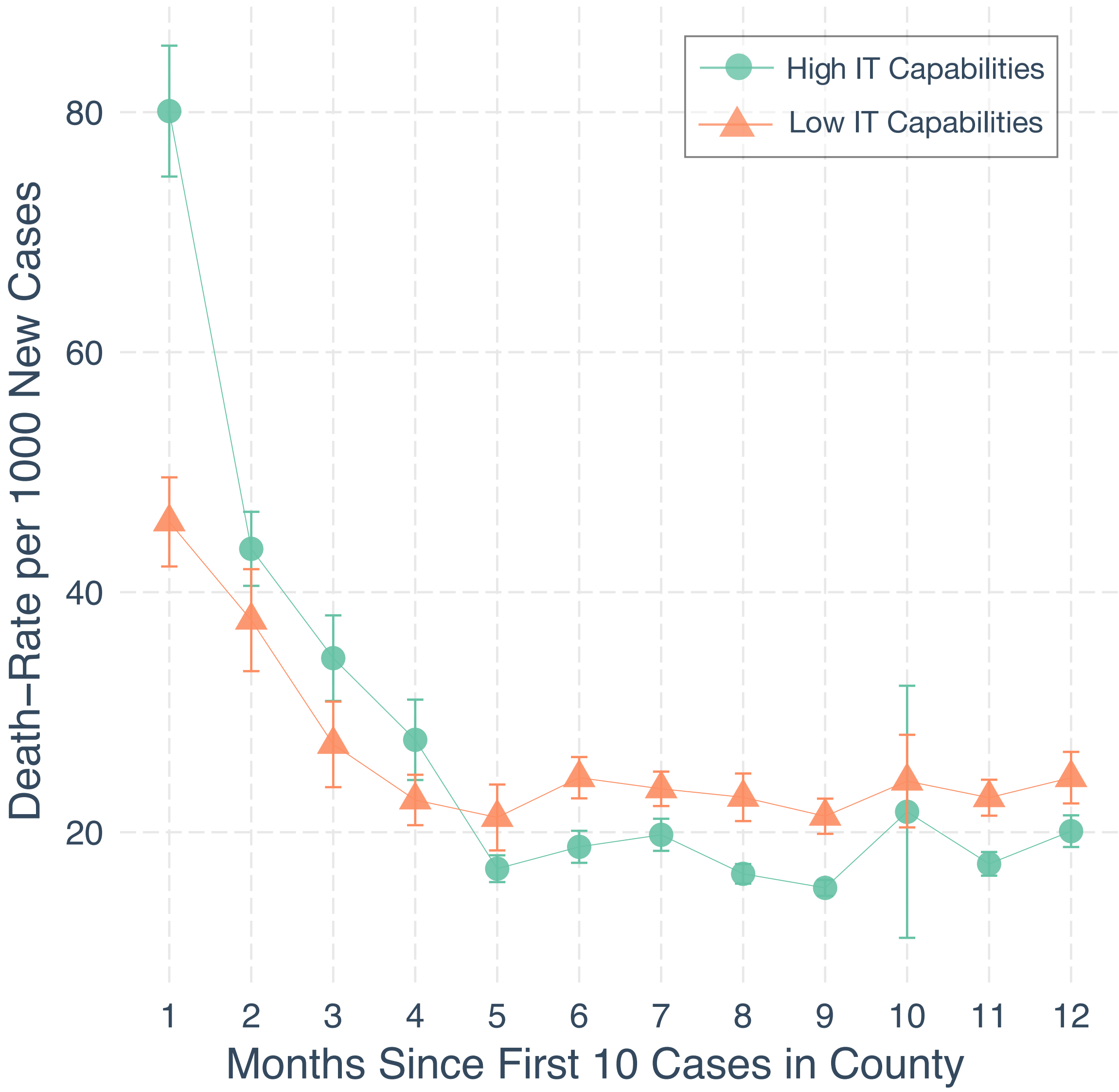

Figure 4 Ratio of new monthly deaths to thousand new monthly cases (lagged 15 days) as a function of the number of months since the first ten cases in the county.

As observed, high IT counties initially fared worse than low IT counties with respect to their death rates per 1000 new COVID-19 cases. However, COVID-19 death-rates for high IT counties decreased at a much faster rate overall than the low IT capability counties. From the first month of the pandemic to the last month of our study period, high IT counties saw their death-rate per thousand new cases decrease from 80.09 (95% CI [74.64,85.54]) to 20.09 (95% CI [18.76,21.40]), whereas low IT counties saw their death rates per capita decrease from 45.85 (95% CI [42.14,49.57]) to 24.54 (95% CI [22.39,26.69]).

[9] We check the results without any lags or lagged one or three weeks. The qualitative results remain unchanged. Robustness checks are available upon request.

To account for the possibility that differences in learning rates between high- and low-IT counties are being driven by systematic differences in covariates, we also present a version of this figure with additional controls for many important covariates. Figure 5 reports predicted deaths rates of counties in different IT-capability bins over time, after controlling for several additional covariates. We report predicted death rates based on linear regressions, with an interaction term between the month and IT bin. In the predictions plotted, all covariates are assumed to take their mean value. Panel A controls for log-population, while panel B includes additional controls such as number of airports, age/gender/race demographic parameters, RUCC urban-rural classification, population and employment densities, internet speeds, and inflow from other counties. We observe that our results in Figure 5 are consistent with our earlier results, even after inclusion of these controls. That is, higher mortality rates in high-IT capability counties in the initial months, followed by lower mortality rates in later months compared to low-IT counties.

To complement these figures, we also present the regression results in Table 4. This table reports OLS estimates of equation (2), based on the number of months since the first ten COVID-19 cases in a county.

$$\text{DeathRate}_{c,t} = \sum_{t=2}^{1} 2\alpha_t \mathbb{1}(M_{c,t} = t) * \text{IT-Index}_c + \gamma_t \mathbb{1}(M_{c,t} = t) + \beta\text{IT-Index}_c + \lambda X_c + \varepsilon_c \tag{2}$$

where $\mathbb{1}(M_{c,t} = t)$ is a dummy for the number of months since the first ten patients in a county, and X_c is a vector of controls and fixed effects.

As observed in Table 4, the coefficients of the interaction of Low health IT index and "Months since 10 cases" all positive and increase over time, across all columns. The change in the magnitude of these coefficients indicates how much faster counties with greater health IT capabilities learned over time. Focusing on the specification in column (3), which includes all controls, we calculate the impact of one standard deviation of the health IT index by adding the direct effect of health IT to the time-varying one, i.e. based on the number of months since ten cases. From the second month of experiencing the pandemic to the twelfth, the gap in mortality between high IT counties and low IT counties expands by 9.4 deaths per thousand cases.

5.2 Failing Faster, Clinical Trials and Long-Run Performance

We also find evidence that learning is related to the presence of COVID-19 related clinical trials of different types at hospitals. We split counties based on the availability of clinical trials, and their type, as shown in Figure 6. We observe that this generates the same pattern of results in terms of initially higher death rates but subsequently lower death rates among high-IT counties. Overall, our results support hypothesis H2 and are consistent with prior observations that healthcare providers in hospitals with clinical trials are likely to experiment with a wide variety of approaches early on, but are quick to learn and adopt best practices based on the knowledge obtained through experimentation and information sharing, in the long run.

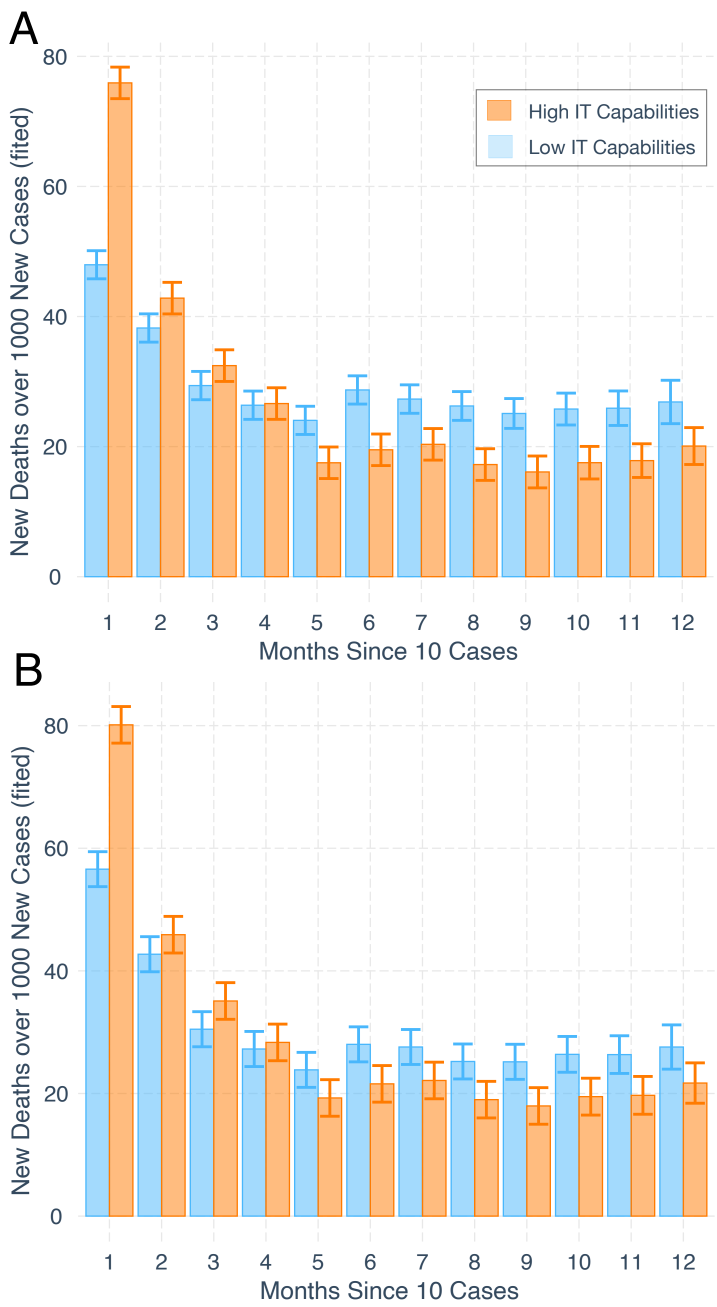

Figure 5 The predicted value of the new monthly deaths over 1000 new -lagged 15 days- monthly cases (along with the 95% confidence intervals) for the 6 months since the first 10 recorded cases for the three groups of counties. A panel includes log-population as control while B panel includes other controls.

A natural next question is whether clinical trials are a complement or substitute for health IT in accelerating learning effects. We are also interested in studying what share of the observed effect of faster learning in high health-IT counties is driven by the presence of clinical trials. Appendix Figure B.2 replicates Figure 4, but restricts our attention to counties without any COVID-19 related clinical trials. We observe a faster rate of learning among high health IT counties, with respect to the decline in mortality rates over time, although the size of this effect is slightly attenuated. This result provides

Table 4 Death Rate by Hospital-IT Intensity and Time

	Dependent variable:		
	New Deaths over 1000 New Cases		
	(1)	(2)	(3)
Months Since 10 Cases (2) × IT_cababilitiesLow	22.599***	20.951***	20.357***
	(4.603)	(4.862)	(4.959)
Months Since 10 Cases (3) × IT_cababilitiesLow	23.862***	20.943***	18.942***
	(4.372)	(4.606)	(4.701)
Months Since 10 Cases (4) × IT_cababilitiesLow	27.987***	23.971***	22.470***
	(4.796)	(4.721)	(5.173)
Months Since 10 Cases (5) × :IT_cababilitiesLow	33.822***	31.658***	28.122***
	(5.037)	(5.565)	(5.676)
Months Since 10 Cases (6) × :IT_cababilitiesLow	36.391***	33.248***	29.988***
	(5.158)	(5.573)	(5.332)
Months Since 10 Cases (7) × :IT_cababilitiesLow	34.249***	33.221***	29.006***
	(5.196)	(5.632)	(5.535)
Months Since 10 Cases (8) × IT_cababilitiesLow	36.009***	33.950***	29.772***
	(4.854)	(5.355)	(5.056)
Months Since 10 Cases (09) × :IT_cababilitiesLow	36.161***	34.372***	30.745***
	(4.689)	(5.064)	(5.059)
Months Since 10 Cases (10) × :IT_cababilitiesLow	35.195***	32.285***	30.469***
	(5.184)	(5.336)	(5.711)
Months Since 10 Cases (11) × IT_cababilitiesLow	35.064***	31.833***	30.311***
	(5.017)	(5.335)	(5.490)
Months Since 10 Cases (12) × IT_cababilitiesLow	33.600***	30.438***	29.728***
	(5.276)	(5.420)	(5.635)
Months Since 10 Cases	yes	yes	yes
IT_cababilitiesLow	−28.286***	−26.282***	−23.349***
	(4.445)	(4.630)	(4.723)
log(Population)	0.376	0.513	1.671**
	(0.667)	(0.653)	(0.811)
Hospital Mortality		0.616**	
		(0.267)	
Median Income			Yes
Change in mobility to workplace (google)			Yes
Inflow from Neigh Counties			Yes
No of Large Airports			Yes
No of Medium Airports			Yes
No of Small Airports			Yes
% of Males			Yes
% of White Pop.			Yes
% of Black Pop.			Yes
% of Indian Pop.			Yes
% of Asian Pop.			Yes
% of 0-25y Pop.			Yes
% of 25-40y Pop.			Yes
% of over 65 Pop.			Yes
Trust in Science			Yes
% of hlds with Broadband			Yes
% of Workers in IT			Yes
Internet Download Speeds			Yes
Employment Density			Yes
Population Density			Yes
No of COVID Clinical Trials			Yes
Observations	26,305	21,421	23,021
Adjusted R^2	0.128	0.135	0.130
Residual Std. Error	38.407	38.439	38.017
Clust. Robust SE	State	State	State
Note:	*p<0.1; **p<0.05; ***p<0.01		

further evidence that the faster rate of learning in high IT counties is only partially explained by a differences in their participation and learning from clinical trials. [10]

[10] In additional analysis, we study whether health IT capabilities and clinical trials are complements with respect to learning effects. We find that while health IT capabilities and clinical trials are commonly co-located, they are actually substitutes for learning best practices over time. In other words, counties with at least moderate levels of health IT capabilities are more likely to conduct clinical trials. Our observations are consistent with health IT playing a role in helping counties learn from clinical trials in neighboring counties, providing evidence of a spillover effect as clinicians learn from their peers' experiences. Results are available upon request.

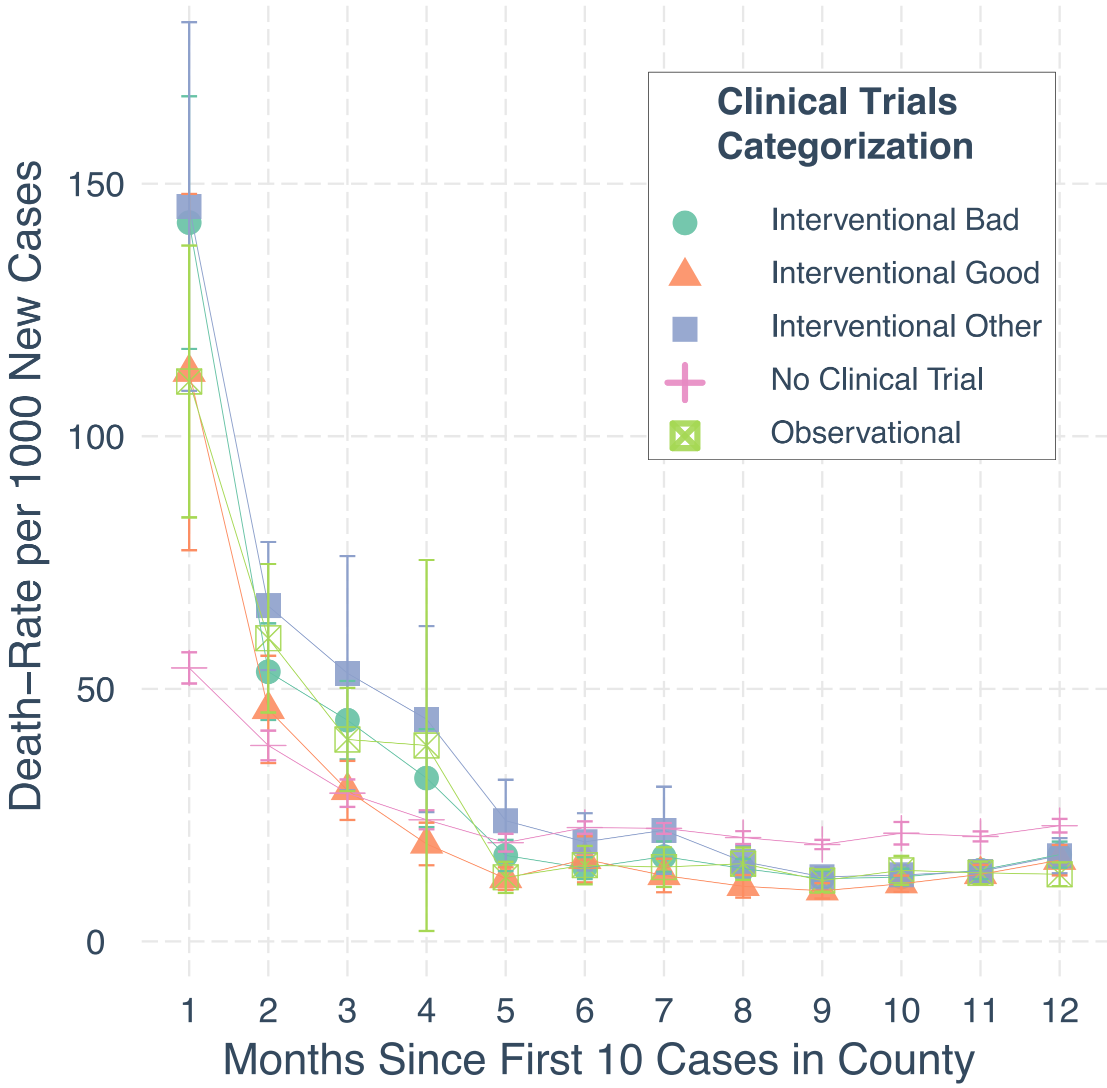

Figure 6 Log-Population weighted mean (and 95% CIs) of the ratio of new monthly deaths to thousand new monthly cases (lagged fifteen days) as a function of the number of months since the first ten cases in the county.

We can study the impact of clinical trials and IT on learning by making two further distinctions. First, we can distinguish between counties with hospitals that pursued clinical trials of drugs that were eventually determined to be unhelpful (e.g. hydroxychloroquine) versus those counties that implemented clinical trials that were eventually determined to be helpful (e.g. Regeneron and other monoclonal antibody treatments). Figure 7 provides point estimates which show that counties with hospitals with both good clinical trials and high IT capabilities performed the best, with respect to cumulative deaths per capita, while counties with hospitals that did not participate in clinical trials and had low health IT capabilities performed the worst. The Figure indicates that counties with high IT capabilities and different types of clinical trial participation exhibited lower cumulative deaths compared to counties with low IT and little or no clinical trial participation. Second, we can distinguish between three types of learning: learning from COVID cases in the same county, learning from COVID cases all over the country/ world and learning from severity of COVID in the county. To do so, we organize counties based on three different notions of time: months since first 10 cases, calendar month and cumulative cases over beds in a county; and plot death rates in Appendix Figure B.4. We find that counties with clinical trials fared better than counties with no clinical trials across all these three types of learning.

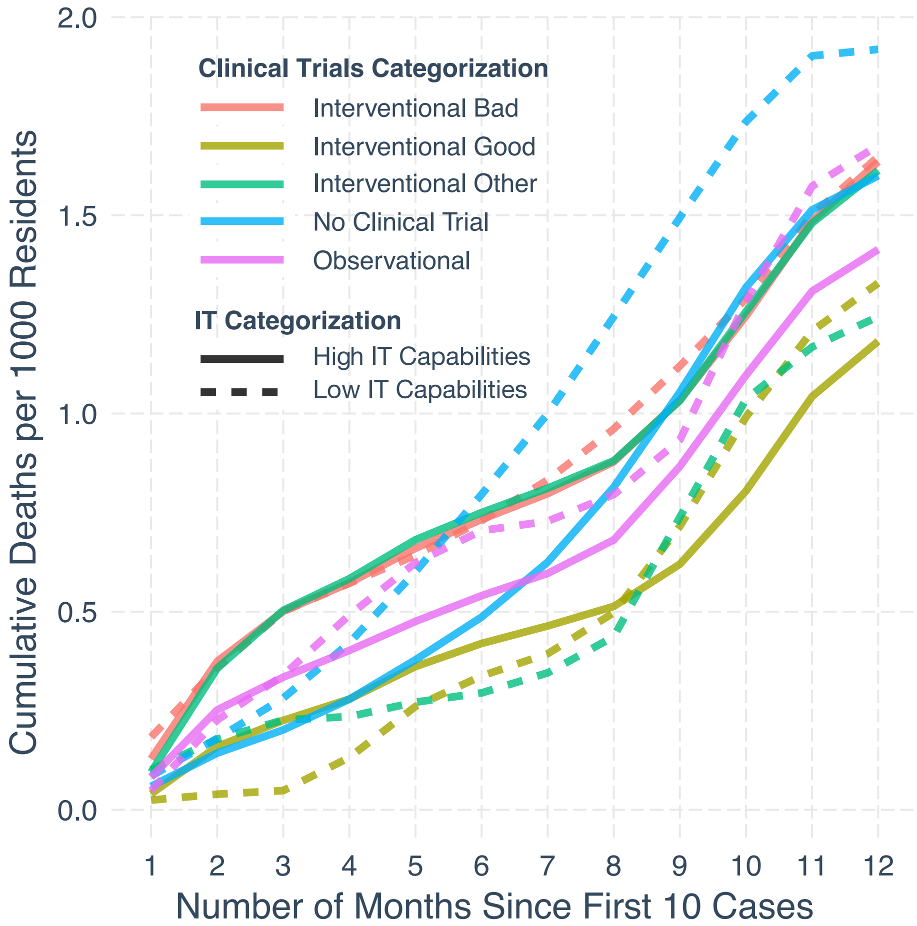

Figure 7 Cumulative deaths per 1000 county residents, by months since first 10 cases. Counties are divided into ten categories based on their IT quality and the type of clinical trial.

5.3 Dry Tinder and Hospital Crowding

Is the faster learning of health IT-intensive counties actually driven by IT, or is this result spurious? This issue is less of a concern in these results than in the previous section focusing on deaths per-capita, because one would need to think of time-varying omitted or endogeneous variables.

However, we can think of two phenomena related to declining death rates that need to be controlled for. The first is a so-called ‘dry-tinder’ effect. This is the idea that some patient populations are more vulnerable to COVID-19. If one county is very homogeneous in its resistance to COVID-19, but another has steep gradations in vulnerability, then we would expect the former to see less of a change in death rates over time. This can be attributed to the maudlin cause that in a population with more heterogeneity in COVID-19 death resistance, only less vulnerable individuals will be left due to differences in population mortality. Hence, we operationalize the concept of dry-tinder by limiting our investigation to counties with only very-high or very-low levels of elderly populations.

Figure B.3 in the Appendix shows that death rates do not seem to be driven by a dry tinder effect. Restricting attention to the youngest (lowest population share below age 65) and oldest (highest population share above 65) decile counties in America shows that low-IT counties performed the worst early on during the pandemic across both county types, but performed slightly better in the oldest counties in the long run. The reduced rate in growth of deaths per capita for low-IT counties in the second panel is consistent with dry-tinder playing a decelerating role for deaths in those counties. Therefore, if dry-tinder were to be driving a spurious relationship between IT and death-rates, we would expect to see the decline in death-rates for IT-intensive hospitals concentrated in the oldest counties. Figure B.3 in the Appendix highlights this analysis and shows the reverse to be true – that the additional decline in death rates for IT-intensive hospitals was concentrated in younger counties, and not present at all in older counties.

A final possibility that we cannot fully exclude is that sicker patients are brought to IT-intensive counties only so long as there has been no large outbreak yet in the high IT county. If the deaths are recorded in the county of death and not the county of residence, this could make the death rate look artificially high in the early months for IT-intensive counties. On a related note, if IT-intensive or clinical-trial using hospitals were more overwhelmed by patients at different points in time, this too could cause their initial poor performance. As mentioned earlier, there is no public data on COVID-deaths at the hospital level. Additionally, there is no public information on crowding at hospitals for the period before July 2020. However, after that period, data has been made available from the US government on bed and ICU capacity (US Government, Department of Health and Human Services 2021). Appendix Figure B.5 reports bed occupancy and ICU occupancy averages for high-IT and low-IT hospitals. We observe that, during the period for which we have data, hospital capacity for the two hospital types closely tracked each other. This partially allays concerns that differential patient mortality rates across counties can be attributed to over-crowding in counties with IT-intensive hospitals in some periods.

5.4 LASSO Analysis

For our final analysis, we use LASSO to determine more precisely how access to IT and clinical trials can help hospitals to learn and improve their treatment of COVID-19 cases for better outcomes. We use LASSO to select parameters, and then run an OLS regression over selected parameters using counties as the unit of analysis. In the first column of Table 5, only the terms associated with learning were selected with LASSO, and all controls were automatically included. In the second column, we conduct the same analysis except the specific controls in the OLS regression model were selected with LASSO as well. LASSO λ factors were chosen to maximize sample fit.

The specific learning parameters we seek to focus on are: Days since first 10 cases – a measure of time since the county encountered COVID; Days since first date – a measure of time since the first cases hit the US; and Cumulative cases per bed – a measure of the amount of experience that a hospital has had with COVID-19 cases. Each of these continuous time dimensions is interacted with two measures of hospital technological capability: health IT-intensity (classified into two bins - low and high) and clinical trial (CT) presence (classified into three bins: bad; neutral or good; and no clinical trials).[11] Table 5 reports the OLS regression on LASSO selected time, IT and CT measures, and controls. The model used is:

$$\begin{aligned} \text{death\_rate} = (&\text{days\_since\_first\_datedays\_since\_10\_cases} + \text{cases\_per\_bed})* \\ &(\text{IT\_capabilities} + \text{CT\_category}) + \text{Controls} \end{aligned} \tag{3}$$

where Controls were not penalized in the Lasso regression in the first column, but are in the second column.

The LASSO analysis suggests a universal effect of slow improvement across all hospitals with calendar time. This is consistent with a global improvement in treatment across all hospitals. Hospitals with low health IT learn from the cumulative number of cases experienced (relative to other hospitals), while they do poorly at retaining this knowledge from their experience over time (see the large negative coefficient on time since first 10 cases interacted with IT). On the other hand, hospitals without any clinical trials also fail to learn from handling large volumes of COVID-19 cases.

Putting both effects together, we develop a better understanding of the disparities in treatment effectiveness of COVID-19 patients across different regions of the country and time. Counties with low health IT hospitals fail to learn best practices from other hospitals over time, or to retain hard-won internal lessons. This is because low-health IT counties lack the necessary infrastructure to share patient health data among healthcare providers, through health information exchanges and other IT-enabled initiatives. Simultaneously, hospitals without clinical trials fail to learn from large case volumes, as they do not benefit from the knowledge gained through experimentation of new types of drugs and treatment procedures. Hence, counties without high health-IT capabilities or experience with clinical trials fall further behind counties comprised of high-IT capability hospitals and/or participate in experimentation through clinical trials. In other words, our LASSO analysis provides support for H3 with respect to the differential effect of learning attributed to health IT capability and clinical trials-based experimentation. Counties with low health IT and lack of clinical trials learn solely based on their recent experience with COVID-19 patients, whereas counties with high IT and participation in clinical trials learn through knowledge sharing from the experience of other providers as well as through their own experience.

6 Discussion

In this study, we provide empirical evidence that hospital implementation and use of health IT systems is associated with reduced mortality attributed to the COVID-19 pandemic. Drawing on county-level data on hospital IT capability, COVID-19 cases, and deaths, we show that counties with greater hospital

[11] We also looked into time since first ten cases for the first hospital group in a county, but there was insufficient variation in this measure to see an effect. Alternatively, it is consistent with that dimension not mattering much for the speed of information diffusion.

Table 5 OLS Regression on LASSO Selected Variables Explaining Death Rate

	Dependent variable:	
	New Deaths over 1000 New Cases	
	(1)	(2)
selected: days since first date of dataset	-0.154^{***}	-0.144^{***}
	(0.001)	(0.001)
selected: days since first 10 cases x IT Low	0.026^{***}	0.024^{***}
	(0.001)	(0.001)
Selected: CT InterventionalGood_mixed_neutral_observational		26.690^{***}
		(0.530)
selected: days since first 10 cases x CT InterventionalGood_mixed_neutral_observational	-0.038^{***}	-0.127^{***}
	(0.001)	(0.002)
selected: cases per available beds x IT Low	-0.027^{***}	-0.024^{***}
	(0.003)	(0.003)
selected: cases per available beds x CT No Clinical Trial	0.060^{***}	0.053^{***}
	(0.003)	(0.003)
Population (log)	3.413^{***}	2.32^{***}
No of Large Airports	0.407	
No of Medium Airports	-1.205^{***}	
No of Small Airports	0.007	
% of Males	-13.628^{***}	-2.69
% of 0-24y Pop.	-46.994^{***}	-29.799^{***}
% of 25-40y Pop.	-131.951^{***}	-145.222^{***}
% of 40-65y Pop.	-39.293^{***}	
RUCC	0.330^{***}	
Employment Density	-0.001^{***}	
% of White Pop	7.397^{***}	-8.806^{***}
% of Black Pop	28.619^{***}	14.791^{***}
% of Indian Pop	21.697^{***}	
% of Asian Pop	22.318^{***}	
Trust in Science	-1.657	
% of IT Workers	-1.055^{***}	
Inflow from Neigh Counties	0.00002^{***}	0.00001^{***}
Population Density	0.002^{***}	0.001^{***}
Internet Download Speeds	-0.009^{***}	
Constant	79.049^{***}	80.930^{***}
N	682,223	682,223
R^2	0.079	0.081
Adjusted R^2	0.079	0.081
Residual Std. Error	52.671	52.612
F Statistic	$2{,}442.414^{***}$	4315^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

IT capability exhibit fewer COVID-19 deaths. Using a battery of controls as well as an instrumental variable analysis, we are able to exclude several non-causal explanations of the focal relationship between health IT and COVID-19 mortality rates, including observable county prosperity, demographics, COVID-19 cases, pre-COVID-19 hospital mortality rates, mobility and pandemic exposure.

Our empirical results suggest that counties with high IT-capability hospitals learned faster how to improve patient treatments and best practices. Further, we observe that hospital participation in clinical trials resulted in learning through experimentation, which is associated with lower mortality rates over time. Although IT-intensive counties were poor performers initially, these counties outperformed their low IT-capability counterparts in the long run, thereby saving lives. Our research suggests that hospital

Table 6 OLS Regressions.

	Dependent variable:	
	New Deaths over 1000 New Cases	
	(1)	(2)
days since 10 cases	-0.131^{***} (0.001)	-0.132^{***} (0.001)
CT Interventional Bad	-2.516^{***} (0.472)	-10.222^{***} (1.080)
CT Interventional Good mixed neutral & observational	-0.771^{**} (0.330)	2.712^{***} (0.672)
High IT	-0.528^{***} (0.146)	-0.509^{***} (0.149)
CT Interventional Bad x High IT		9.292^{***} (1.191)
CT Interventional Good mixed neutral & observational x High IT		-4.312^{***} (0.738)
Controls		
Population (log)	4.351^{***} (0.127)	4.354^{***} (0.127)
No of Large Airports	-1.097^{***} (0.254)	-1.027^{***} (0.255)
No of Medium Airports	-1.600^{***} (0.138)	-1.628^{***} (0.138)
No of Small Airports	0.044^{***} (0.015)	0.039^{***} (0.015)
% of Males	-7.758 (4.719)	-9.272^{**} (4.721)
% of 0-24y Pop.	-33.729^{***} (2.533)	-34.099^{***} (2.534)
% of 25-40y Pop.	-125.909^{***} (3.917)	-125.119^{***} (3.919)
% of 40-65y Pop.	-9.829^{**} (4.147)	-10.037^{**} (4.147)
RUCC	0.028 (0.046)	0.042 (0.046)
Employment Density	-0.0004^{***} (0.0001)	-0.0005^{***} (0.0001)
% of White Pop.	1.576 (1.873)	1.870 (1.873)
% of Black Pop.	28.548^{***} (1.892)	28.822^{***} (1.892)
% of Indian Pop.	18.406^{***} (2.391)	18.729^{***} (2.391)
% of Asian Pop.	11.411^{***} (4.092)	10.924^{***} (4.093)
Trust in Science	-2.529^{**} (1.168)	-2.260^{*} (1.168)
% of IT Workers	-1.341^{***} (0.113)	-1.333^{***} (0.113)
Inflow from Neigh Counties	0.00001^{***} (0.00000)	0.00001^{***} (0.00000)
Population Density	0.001^{***} (0.0001)	0.001^{***} (0.0001)
Internet Download Speed	-0.006^{***} (0.002)	-0.007^{***} (0.002)
Constant	45.563^{***} (4.055)	45.806^{***} (4.055)
Observations	682,223	682,223
R^2	0.071	0.071
Adjusted R^2	0.071	0.071
Residual Std. Error	52.897	52.894
F Statistic	$2{,}273.552^{***}$	$2{,}095.853^{***}$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

health IT systems allow counties to develop greater digital capabilities related to experimentation and knowledge sharing. In turn, these organizational capabilities enable clinicians to better understand and improve best practices for treatment of COVID-19 patients, gather real-time information on patients' health status and test results, and share patient information to disease surveillance registries, thereby reducing COVID-19 mortality rates.

Our research is not only one of the first studies to document the impact of health IT capabilities on health outcomes during a pandemic outbreak, but also among the early studies to highlight the role of IT-enabled learning and experimentation effects on organizational performance. Specifically, with respect to causal mechanisms, our results show that high health IT counties have a much higher rate of clinical trials, which in turn may contribute to improvements in their rate of learning over time. In other words, information transfer and exchange of best practices for treatment of COVID-19 patients, based on data from clinical trials, may contribute to the significant reduction in observed mortality rates within high IT counties. From an operations lens, our research indicates that counties can benefit by improving information sharing across healthcare providers and make relevant investments in clinical trials, in order to improve learning and reduce mortality rates. In a LASSO analysis, we provide further evidence for our findings that IT and CTs are important for learning. LASSO models consistently selected variables which show that hospitals without IT struggle to learn except from "learning by doing" (demonstrating a role for IT in information diffusion), and hospitals without clinical trials fail to learn from the knowledge obtained through experimentation.

What is the size of this effect? Our estimation results indicate that, after flexibly controlling for when a county was first exposed to COVID-19 and a state-level fixed effect, one standard deviation increase in our county-level health IT index is associated with .063 fewer deaths per thousand residents. For a US population of 331 million, this suggests that a one standard deviation increase in county-level, health IT capability across all hospitals would have reduced COVID deaths by 20,853 in the United States through April 5, 2021. This estimate represents a substantial reduction in mortality rates that may have been avoided if hospitals and counties had invested in relevant health IT capabilities to improve their digital capabilities and support dissemination of public health information and best practices to clinicians involved in treating COVID-19 patients.

Our study has several implications for research and practice. While the body of knowledge on best practices for treatment of COVID-19 patients has evolved in the relatively short period of time since the advent of the pandemic, scant attention has been paid to the impact of health IT on the performance of healthcare provider organizations. Such IT-enabled capabilities provide clinicians with real-time information on patient health data and test results, as well as clinical decision support supported by timely knowledge of best practices, and are likely to be useful in reducing the spread of pandemic outbreaks, resulting in lower mortality rates. Although the theory of organizational learning (Argyris 1977) has been proposed to study the role of learning in the implementation of information systems, very little is known about how health IT can contribute toward dissemination of real-time knowledge in the context of a pandemic outbreak or treatment of other infectious diseases.

Our research represents an early effort to explore the impact of different types of health IT systems on information management and knowledge dissemination during the COVID-19 pandemic, and how they contribute to organizational learning to improve delivery of patient care across time. From an operations perspective, our findings suggest that counties should invest in more resources to improve their health IT infrastructure as well as knowledge dissemination from participation in clinical trials. In this respect, our empirical research suggest the presence of single- and double-loop learning effects as clinicians not only absorb best practices learned from others, but also rely on clinical trials to question conventional wisdom and refine their understanding of best practices for treatment of COVID-19 patients. These learning effects can be attributed to hospital participation in clinical trials which enabled them to reduce COVID-19 death rates over time through a process of experimentation with different types of treatments.

Our research is not without limitations. A major limitation is that we only have lagged archival information about hospital health IT that is reported annually, which precludes us from studying the impact of changes in IT use at a granular level (such as weekly or monthly time windows). Further, our measure of the county-level health IT index is based solely on hospital adoption and use of health IT, and does not include use of health IT in outpatient clinics, nursing homes, and other tertiary care centers, which is not easily available in research databases such as HIMSS or AHA IT Supplement. Finally, we are also limited by COVID-19 death and case information being available only at the county level, which averages out variations in heterogeneity across many hospitals in populous counties like New York County (i.e. Manhattan). Nevertheless, our research represents one of the first attempts to study the impact of hospital IT capabilities on county-level mortality rates of COVID-19 patients across the United States. Furthermore, it provides a useful lens to explore the role of organizational learning and experimentation through hospital participation in clinical trials and their impact on patient health outcomes.

7 Conclusion

We study the impact of health IT on COVID-19 mortality rates across the United States, based on county-level data on hospital implementation of IT systems. Specifically, we focus on the role of organizational learning using clinical trials as an important mechanism through which learning effects may be manifested in terms of patient health outcomes. Our empirical analyses are based on a rich data set of COVID-19 cases and mortality rates across the United States, and the corresponding use of health IT systems by hospitals measured at the county level. We find that health IT adoption and use reduced patient mortality rates by improving the rate of learning in counties that exhibit higher levels of health IT capabilities. We also show that high health-IT counties are able to learn much faster than low health-IT counties over time, which results in significant reduction in mortality rates compared to low-IT counties. Our results also indicate that counties with greater hospital participation in clinical trials also performed better than counties with little or no clinical trial participation. Altogether, our research represents one of the first studies to provide empirical evidence on the impact of health IT and clinical trials on patient health outcomes during the COVID-19 pandemic, while focusing on the role of organizational learning effects through two mechanisms - experimentation and knowledge sharing.

References

Amarasingham, R., L. Plantinga, M. Diener-West, D. J. Gaskin, N. R. Powe. 2009. Clinical information technologies and inpatient outcomes: a multiple hospital study. *Archives of internal medicine*, 169 (2), 108-114.

American Journal of Managed Care, Staff of. 2020. A timeline of covid-19 developments in 2020. URL https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020.

Argyris, C. 1977. Organizational learning and management information systems. *Accounting, Organizations and Society*, 2 (2), 113-123.

Atasoy, H., P.-y. Chen, K. Ganju. 2018. The spillover effects of health it investments on regional healthcare costs. *Management Science*, 64 (6), 2515-2534.

Avgerinos, E., B. Gokpinar. 2018. Task variety in professional service work: When it helps and when it hurts. *Production and Operations Management*, 27 (7), 1368-1389.

Ayabakan, S., I. Bardhan, Z. Zheng, K. Kirksey. 2017. The impact of health information sharing on duplicate testing. *MIS Quarterly*, 41 (4),.

Ayabakan, S., I. R. Bardhan, Z. Zheng. 2021. Impact of telehealth on healthcare resource utilization: A quasi-experimental patient-level study. *Working paper*, .

Bayo-Moriones, A., M. Billon, F. Lera-López. 2017. Are new work practices applied together with ict and amt? *The International Journal of Human Resource Management*, 28 (4), 553-580.

Benkard, C. L. 2000. Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90 (4), 1034-1054.

Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, et al. 2020. How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data, .

ClinicalTrials.gov. 2021. Background information about clinical trial online registry. URL https://clinicaltrials.gov/ct2/about-site/background.

Collis, A., K. Garimella, A. Moehring, M. A. Rahimian, S. Babalola, D. Shattuck, D. Eckles, S. Aral. 2020. Global survey on covid-19 beliefs, behaviors, and norms. *MIT Technical Report*, .

Dong, E., H. Du, L. Gardner. 2020. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20 (5), 533-534.

Dranove, D., C. Forman, A. Goldfarb, S. Greenstein. 2014. The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, 6 (4), 239-70.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, et al. 2020. Estimating the effects of non-pharmaceutical interventions on covid-19 in europe. *Nature*, 584 (7820), 257-261.

Huckman, R. S., G. P. Pisano. 2006. The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science*, 52 (4), 473-488.

Janakiraman, R., E. Park, E. Demirezen, S. Kumar. 2021. The effects of health information exchange access on healthcare quality and efficiency: An empirical investigation. *Management Science*, Forthcoming.

Kohli, R., S. S.-L. Tan. 2016. Electronic health records: how can is researchers contribute to transforming healthcare? *Mis Quarterly*, 40 (3), 553-573.

Kraemer, M. U., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. Du Plessis, N. R. Faria, R. Li, W. P. Hanage, et al. 2020. The effect of human mobility and control measures on the covid-19 epidemic in china. *Science*, 368 (6490), 493-497.

Kwon, J., M. E. Johnson. 2014. Proactive versus reactive security investments in the healthcare sector. *Mis Quarterly*, 38 (2), 451-A3.

Ledford, H. 2020. Why do covid death rates seem to be falling? *Nature*, 587 190-192. URL https://www.nature.com/articles/d41586-020-03132-4.

Lee, J., J. S. McCullough, R. J. Town. 2013. The impact of health information technology on hospital productivity. *The RAND Journal of Economics*, 44 (3), 545-568.

Loewenstein, G., J. S. Lerner. 2003. The role of affect in decision making., .

Nelson, R. R., S. G. Winter. 1982. The schumpeterian tradeoff revisited. *The American Economic Review*, 72 (1), 114-132.

Pisano, G. P., R. M. Bohmer, A. C. Edmondson. 2001. Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management science*, 47 (6), 752-768.

Self, W. H., M. W. Semler, L. M. Leither, J. D. Casey, D. C. Angus, R. G. Brower, S. Y. Chang, S. P. Collins, J. C. Eppensteiner, M. R. Filbin, et al. 2020. Effect of hydroxychloroquine on clinical status at 14 days in hospitalized patients with covid-19: a randomized clinical trial. *Jama*, 324 (21), 2165-2176.

Sharma, L., A. Chandrasekaran, K. K. Boyer, C. M. McDermott. 2016. The impact of health information technology bundles on hospital performance: An econometric study. *Journal of Operations Management*, 41 25-41.

Sosa, J. A., H. M. Bowman, T. A. Gordon, E. B. Bass, C. J. Yeo, K. D. Lillemoe, H. A. Pitt, J. M. Tielsch, J. L. Cameron. 1998. Importance of hospital volume in the overall management of pancreatic cancer. *Annals of surgery*, 228 (3), 429.

Tucker, A. L., I. M. Nembhard, A. C. Edmondson. 2007. Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management science*, 53 (6), 894-907.

US Government, Department of Health and Human Services. 2021. Covid-19 reported patient impact and hospital capacity by state - archive repository. URL https://healthdata.gov/dataset/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/4cnb-m4rz.

Von Hippel, E., M. J. Tyre. 1995. How learning by doing is done: problem identification in novel process equipment. *Research policy*, 24 (1), 1-12.

Wani, D., M. Malhotra. 2018. Does the meaningful use of electronic health records improve patient outcomes? *Journal of Operations Management*, 60 1-18.

Xiong, C., S. Hu, M. Yang, W. Luo, L. Zhang. 2020. Mobile device data reveal the dynamics in a positive relationship between human mobility and covid-19 infections. *Proceedings of the National Academy of Sciences*, 117 (44), 27087-27089.

Zhang, L., S. Ghader, M. L. Pack, C. Xiong, A. Darzi, M. Yang, Q. Sun, A. Kabiri, S. Hu. 2020. An interactive covid-19 mobility impact and social distancing analysis platform. *medRxiv*, .

Zollo, M., S. G. Winter. 2002. Deliberate learning and the evolution of dynamic capabilities. *Organization science*, 13 (3), 339-351.